Evolution in Northern Neotropical Catfishes – Integrating Genomics and Morphometrics

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde (Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Madlen Stange

aus

Deutschland

Promotionskommission

Prof. Dr. Marcelo Sánchez-Villagra (Vorsitz und Leitung der Dissertation)

Prof. Dr. Walter Salzburger

Prof. Dr. Lukas Keller

Dr. Juan Ignacio Montoya-Burgos (Gutachter)

Zürich, 2017

Dissertation

Evolution in Northern Neotropical Catfishes – Integrating Genomics and Morphometrics

Content

Acknowledgements

This dissertation would not have been written if it were not for the curiosity of my supervisor Marcelo R. Sánchez-Villagra. Thankfully, he got bored (just a little) with terrestrial vertebrates and became curious about aquatic species diversity. I was lucky enough that he hired me to do the job. I thank Marcelo for his intellectual guidance and financial support as well as his efforts to advance women in science.

I want to thank Sandra Schwarte for teaching me how to work thoroughly, thoughtful, and conceptual. I am so glad for having been your student assistant and your friend. You made me the scientist I am today. I thank Michael Lenhard for being the most inspiring and passionate professor I have met during my undergraduate and graduate studies. Thank you for actually caring for your students. I thank Stefanie Hartmann and Detlef Groth for realizing how important R will be for any biologist and for teaching it so well and patiently to us ignorant. I thank all the heads and members of the different working groups I had the honor to be a temporary part of, for the inspiring discussions, support, and fun obviously: the group of Ralph Tiedemann (Uni Potsdam), Michael Lehnhard (Uni Potsdam), Walter Salzburger (Uni Basel), and Marcelo R. Sánchez-Villagra (UZH).

I want to thank Walter for allowing me to do my Master thesis at his lab. It started a series of events for which I am very grateful today.

Huge thanks go to my many co-authors and collaborators from Switzerland, Panama, and Venezuela, in particular Jorge Domingo Carrillo Briceño (UZH), Richard G. Cooke from the Smithsonian Tropical Research Institute, and Cathy Villalba and Gloria Sánchez from Venezuela. Fieldwork requires the help of many and I want to state that I would be nowhere close the achieved sample size without the support of the many helpers. Finally, I thank my family and friends for their support during this very long educational process: Stefan, Alex, Laura, and Astrid, you are the best!

Summary

Transitions between the marine and freshwater macrohabitat have occurred repeatedly in the evolution of teleost fishes. In this PhD thesis five genera of northern Neotropical sea catfish (Ariidae), a marine teleost lineage that repeatedly adapted to freshwater habitats, was investigated with regard to shape changes in relation to macrohabitat transition. An underlying fundamental assumption is that many morphological changes are adaptive in nature and it was examined here whether similar morphological changes occurred with each freshwater transition. Two skeletal features, the opercle bone, that has been shown to be a good marker for habitat adaptation in teleost fishes, and the skull, a commonly used marker for habitat adaptation in terrestrial vertebrates, were investigated applying 2D and 3D geometric morphometrics, respectively. The resulting shape data was analysed using phylogenetically informed multivariate statistics. The application of a mitochondrial DNA marker helped to verify species identities.

The distribution of the investigated group makes it a rich subject to address the fundamental question on faunal evolution in the northern Neotropics. Another section of this thesis concerns the investigation of the dating of lineage splitting events among Caribbean and Eastern Pacific ariid species, solving questions on the taxonomy and phylogeny of the group but also testing conflicting hypotheses on the timing of the rise of the Panamanian Isthmus. To do so, a novel approach using Bayesian divergence time estimation was developed that constructs time-calibrated phylogenetic hypotheses from single-nucleotide polymorphisms and fossils for internal-node calibration. In the course of this study new species and freshwater populations of brackish/marine species were found and partially validated by mtDNA, nuclear DNA, or shape analysis, or all of the above together. Comprehensive morphometric analyses showed that habitat has a larger effect on opercle shape than on skull shape, but phylogenetic relationships largely determine the position of species in morphospace for both features. Freshwater species occupy a smaller area in morphospace than brackish and marine species. The Bayesian divergence time estimation revealed that most lineage splitting events among species from the two oceans happened during the Miocene, thus supporting findings for a longer and more complex formation history of the Panamanian Isthmus.

Zusammenfassung

Während der Evolution der echten Knochenfische (Teleostei) kam es häufig zu Transitionen zwischen den Makrohabitaten Salz- und Süßwasser. In der vorliegenden Doktorarbeit wurden fünf Gattungen der im salzwasserlebenden Ariidae (Kreuzwelse) aus den nördlichen Neotropen untersucht. Sie zeichnen sich dadurch aus, dass Mitglieder der einzelnen Gattungen sich wiederholt an Süßwasserhabitate angepasst haben. Die zugrundeliegende Hypothese ist, dass a) Formveränderungen der jeweiligen Art auf Grund von Anpassung an das neue Habitat auftreten können, und b) im gleichen Habitat die immer gleiche Form begünstigt werden könnte. Zwei Kapitel der vorliegenden Arbeit befassen sich mit den Formveränderungen, die mit dem Habitatwechsel einhergehen. Dazu wurden zwei Skelettmerkmale untersucht. Zum einen der Kiemendeckel, der bereits in anderen Knochenfischen als Marker für Makrohabitattransitionen eingesetzt wurde. Zum anderen der Schädel, der besonders bei terrestrischen Tieren Anwendung bei der Bestimmung von Habitattransitionen findet. Dazu wurde 2D respektive 3D geometrische Morphometrie angewandt und mit Hilfe von phylogenetisch-gestützter Statistik analysiert.

Das Verbreitungsgebiet der untersuchten Arten eignet sich hervorragend um fundamentale Fragen der tierischen Evolution in den nördlichen Neotropen zu untersuchen. Ein weiteres Kapitel beschäftigt sich der zeitliche Datierung von Artaufspaltungen der Karibik lebenden Arten von denen im östlichen Pazifik lebenden Arten. Diese kann Aufschluss über die Taxonomie und Phylogenie der Gruppe geben, sowie zu der Diskussion beitragen, ob der Isthmus von Panama nicht eventuell älter ist als bisher allgemeinhin angenommen wurde. Zunächst wurden die Artzugehörigkeiten aller untersuchten Individuen mittels mitochondrialen DNS Marker verifiziert. Die Formenanalyse ergab, dass es habitatspezifische Kiemendeckelformen gibt. Allerdings hat der Verwandtschaftsgrad einen stärkeren Einfluss auf Kiemendeckel- und Schädelform als das jeweilige Habitat aus denen die Fische stammten. Generell lassen sich keine Hinweise auf Konvergenz von Schädelformen von Arten aus dem gleichen Habitat finden. Der Formenraum, den Süßwasserarten einnehmen, ist viel kleiner als der von Salz- und Brackwasserlebenden Arten. In dieser Arbeit wurden neue Arten und Süßwasserpopulationen von Salz- und Brackwasserlebenden Arten gefunden und teilweise mittels mitochondrialer, nuklearer DNS oder Formenanalyse validiert. Die zeitliche Einordnung der Artaufspaltungen über die Entstehungsgeschichte der heutigen hier untersuchten Ariidae der nördlichen Neotropen unterstützt eine frühere Erhebung des Isthmus von Panama, und zwar bereits vor etwa 10 Millionen Jahren.

Introduction

Evolution in Northern Neotropical Catfishes – Integrating Genomics and Morphometrics

Personal motivation

Almost wherever we go in the natural world, whatever we lay our eyes upon, a spectacular and stunning diversity of organisms surrounds us. An innate curiosity makes me wonder how it came to such a variety that is expressed in different shapes, sizes, colours, physiologies, and more. As an evolutionary biologist, I pose myself a question that already occupied Charles Darwin:

"How have all those exquisite adaptations of one part of the organisation to another part, and to the conditions of life, and of one distinct organic being to another being, been perfected?" – (On the Origin of Species by Means of Natural Selection, 1859, p.39, Dover thrift edition)

Selectional (sexual and natural) forces and random drift are drivers of diversification of populations in natural habitats. The specific evolutionary forces of phenotypic and genotypic diversification that promote evolution in general and more specifically adaptation and speciation can exhibit the same patterns and results from the same processes or not. To understand global patterns and processes of evolution, adaptation, and speciation, it is imperative to understand diversification on the smallest scale, that is closely related populations and species.

Recently, I was once more reassured that focusing on a specific clade can be most rewarding, as I listened to an inspiring talk by Prof. Erika J. Edwards during the 2016 Evolution Meeting in Texas, USA. She talked about the relevance of studying individual taxa, instead of big phylogenies and the major patterns of evolution. To make her point she presented the case study on the plant *Viburnum*, about which she and her colleagues study the evolution of leaf form and function. They found the cause for serration in leaves from temperate taxa versus entire leaves found in tropical taxa. To make a long story short, this is because temperate taxa form their leaves in the resting stage of the bud. As a consequence the individual leaf has to be folded up in the most practical way, the result being a serrated leaf. Tropical leaves grow as they emerge; they are not formed within buds, because leaf growth never ceases. The pattern of serrated versus entire leaves was obvious, but only after having been able to put the phylogeny together. The study of the process could only follow after the pattern was identified. Only the relentless, curiosity-driven study of a clade that has an excellent distribution pattern and a well-defined phylogeny made the unravelling of the process possible.

In my PhD thesis I focused on a group of teleost fishes with an elevated rate for speciation and morphological evolution (Rabosky *et al.*, 2013), namely Ariidae (sea catfishes). As the curious scientist I am asking: *"Why is this?"* Ariidae feature a special ability and that is the propensity to adapt to freshwater habitat and to speciate. During their evolutionary history they repeatedly occupied freshwater habitats (Marceniuk & Menezes, 2007), and as we saw during our fieldwork, they are still doing so. This ability could be a potential driver of diversification therefore I aimed to investigate

the effect of habitat on the shape evolution in sea catfishes as means of a first approximation to understand the evolution of this understudied radiation. To understand the evolution of this group the geographic component is very important. In the case of Ariidae, their evolution is closely tied to a major event at a global level, namely the origin of the connection of North and South America in geological time scales. This event greatly affected the oceans worldwide and the evolution of Ariidae, given the separation of two oceans and the creation of new habitats in the newly formed Caribbean Sea.

General information about Ariidae (sea catfishes)

Ariidae are distributed in all tropical and sub-tropical regions worldwide in coastal and near-coastal areas and rivers (Marceniuk & Menezes, 2007). It is hypothesised that they radiated from the Tethys Sea in southern Gondwanaland, at a time when South America and Africa were still connected but Australia was already disconnected from Gondwanaland (Betancur-R. *et al.*, 2007). The systematics of the Ariidae was recently revised (Acero P. & Betancur-R., 2007; Betancur-R. *et al.*, 2007, 2012; Marceniuk & Menezes, 2007; Betancur-R., 2009; Marceniuk, Menezes, *et al.*, 2012) either based on morphology (Acero, and Marceniuk and colleagues) or by a combination of molecular and morphological markers (Betancur-R. and colleagues). Molecular studies were based on few mitochondrial and nuclear genes. Those studies resulted in the revision of the genus *Arius* that formerly contained 121 species (Teugels, 1996) spread on several continents. As a consequence of the revisions, *Arius* is now restricted to Old World taxa, and New World *Arius* taxa were redefined into *Notarius*, *Cathorops*, *Sciades*, *Potamarius*, *Occidentarius*, *Genidens*, *Bagre,* and *Ariopsis* counting some 60 species in the Americas (Froese & Pauly, 2016). The presented chapters focus on the Ariidae from the northern Neotropics, a region that encompasses northern South-America and Meso-America. Studying Ariidae of that region has the advantage that we can make use of a growing Neogene fossil record that includes otoliths, partial skulls, and spines, which are useful for time calibration of molecular trees.

New species are frequently described, still, from Meso- and South America (Acero P. & Betancur-R., 2002; Betancur-R. & Acero, 2004; Betancur-R. & Acero P., 2005; Marceniuk, 2007; Betancur-R. *et al.*, 2008; Marceniuk, Betancur-R., *et al.*, 2012). New discoveries, as presented in this PhD thesis as well as in the work of other researchers, lead to more and more new species in need of formal description (Betancur-R., 2009; da Silva *et al.*, 2016; Stange *et al.*, 2016). Due to this fact I aimed to infer a solid phylogenetic hypothesis for the individuals that were studied in exactly this PhD thesis rather than assuming their relationship based on previous studies. In collaboration with Dr. Michael Matschiner, we present a new phylogenetic hypothesis in Chapter 2, which is used in Chapter 3 to investigate shape evolution in a of phylogenetic framework.

Relevance of the geographic setting of the sampling sites

Ariidae of the northern Neotropics were probably strongly affected by the uplift of the Panamanian Isthmus. The Isthmus of Panama connects South and North America and divides the Western Atlantic and Eastern Pacific, permanently since 2.8 million years ago (Ma) (O'Dea *et al.*, 2016). This complete schism resulted in tremendous oceanic, climatic, and biotic changes, and the geographic separation of formerly genetically connected marine populations and simultaneously the connection of terrestrial realms. A significant taxonomic terrestrial exchange took place from 2.6 Ma on and is known as the Great American Biotic Interchange (GABI), verified by numerous fossils and molecular studies (Woodburne, 2010). Currently, it is debated whether the land bridge might have formed an earlier permanent connection than commonly assumed. Recent studies argue that the Panamanian Arc initially collided with South America 25-23 Ma (Farris *et al.*, 2011) and that the Central American Seaway (CAS) closed as early as the Middle Miocene (13-15 Ma) (Montes *et al.*, 2012, 2015). As a consequence, terrestrial taxa in North and South America should have had the potential to migrate in both directions. The closed CAS supposedly lead to pulses of dispersal of land animals ~ 23 and ~ 7 Ma and to dispersal barriers for marine organisms ~ 8 Ma (Bacon, Silvestro, *et al.*, 2015; Bacon, Silvetsro, *et al.*, 2015; Carrillo *et al.*, 2015). But the observed dispersal rates were much lower than what was observed after 2.6 Ma what is classically called the GABI. Reasons for a delayed taxonomic exchange of North and South American lineages other than an open CAS was argued to be a lack of suitable habitats, that only emerged after the onset of glaciation 3.3 Ma (Bacon *et al.*, 2016; Bloch *et al.*, 2016). An even more recent review on the geological, paleontological, oceanographic, and molecular evidences agreed on the initial collision but found no evidence for the closure of the CAS before the Pliocene and argued for a formation of the Isthmus *sensu stricto* at 2.8 Ma (O'Dea *et al.*, 2016).

There is abundant evidence of dispersal of terrestrial animals and plants over water in pre-GABI times that is in many cases explained by "rafting" on floating islands (de Queiroz, 2014). And it is a plausible explanation where uncertainty of geographic connections prevails. In order to test true vicariance, as in the case of the Central American Seaway and the Panamanian Isthmus, marine organisms provide more reliable results. The evolution of many clades found in coastal marine areas in the Tropical Eastern Pacific and Western Atlantic has been greatly influenced by the rise of the Isthmus of Panama, as echinoderms (Lessios, 1981), snapping shrimps (*Alpheus*) (Knowlton *et al.*, 1993), angel sharks (*Squatina*) (Acero P. *et al.*, 2016), or bonnet head sharks (*Sphyrna tiburo*) (Martin *et al.*, 1992).

Since the Ariidae are nowadays found in the Western Atlantic and Eastern Pacific and diversified well before the formation of Isthmus of Panama, we investigated their lineage splitting patterns. In Chapter 2 we draw conclusions

about their ancestral distribution area as well as about the timing of lineage splitting events in relation to the ongoing discussion about the formation of the Panamanian isthmus.

Macrohabitat transitions as driver of speciation

Macrohabitat transitions in fishes can lead to the evolution of new species via divergent selection and reproductive isolation. Transitions that promoted speciation include transitions from the bottom-associated habitat to the open-water column, and within the water-column, diversification along the depth-gradient. Examples for this are Antarctic notothenioid fishes (Matschiner *et al.*, 2011; Rutschmann *et al.*, 2011), Nicaraguan Crater Lake cichlids (Barluenga & Meyer, 2010), as well as African lake cichlids (depth-gradient) (Seehausen *et al.*, 2008). Lake-stream transitions repeatedly promoted the diversification of sticklebacks (Berner *et al.*, 2009). And last but not least, the colonization of freshwater habitats from marine ancestors (and vice versa) is a common scenario in various teleost families including sticklebacks (Hohenlohe *et al.*, 2010), needlefish (Lovejoy & Collette, 2001), Neotropical silversides (Bloom *et al.*, 2013), terapontid grunters (Davis *et al.*, 2012) and sea catfishes (Betancur-R., 2010). Such transitions from marine to freshwater macrohabitats implicate varied adaptations to different environmental conditions in, e.g., salinity, pH, flow conditions, oxygen contents, or feeding ecology (type of pray and way of feeding). In the course of adaptation to these new environments morphological and genetic changes are to be expected. The evolutionary history of habitat transitions from the marine back to freshwater habitat, with the availability of intermediate species with brackish occurrence makes the Ariidae a good system to study marine-freshwater transitions and their patterns. In Chapter 1 the shape evolution of the opercle bone, a facial bone covering the gills, was investigated in relation to habitat. This bone has been shown to be a good marker for habitat adaptation in sticklebacks (Kimmel *et al.*, 2012), cichlids (Wilson *et al.*, 2015), and Antarctic notothenioids (Wilson *et al.*, 2013). In Chapter 3, I went on and looked at the shape changes in the ariid skull, a composite structure made of several bones, holding visual and olfactory systems, as well as tooth plates. The basal expectation is that the more complex the investigated structure is, the more evident should be any sign of adaptation.

References

Acero P., A. & Betancur-R., R. 2002. *Arius cookei*, a new species of ariid catfish from the tropical American Pacific. *Aqua, J. Ichthyol. Aquat. Biol.* 5: 133–138.

Acero P., A. & Betancur-R., R. 2007. Monophyly, affinities, and subfamilial clades of sea catfishes (Siluriformes: Ariidae). *Ichthyol. Explor. Freshwaters* 18: 133–143.

Acero P., A., Tavera, J.J., Anguila, R. & Hernández, L. 2016. A New Southern Caribbean Species of Angel Shark (Chondrichthyes, Squaliformes, Squatinidae), Including Phylogeny and Tempo of Diversification of American Species. *Copeia* 104: 577–585.

Bacon, C.D., Molnar, P., Antonelli, A., Crawford, A.J., Montes, C. & Vallejo-Pareja, M.C. 2016. Quaternary glaciation and the Great American Biotic Interchange. *Geology* 44: 375–379.

Bacon, C.D., Silvestro, D., Jaramillo, C., Smith, B.T., Chakrabarty, P. & Antonelli, A. 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. *Proc. Natl. Acad. Sci.* 112: 6110–6115.

Bacon, C.D., Silvetsro, D., Jaramillo, C., Smith, B.T., Chakrabarty, P. & Antonelli, A. 2015. Reply to Lessios and Marko et al .: Early and progressive migration across the Isthmus of Panama is robust to missing data and biases. *Proc. Natl. Acad. Sci. Lett.* 112: E5767-5768.

Barluenga, M. & Meyer, A. 2010. Phylogeography, colonization and population history of the Midas cichlid species complex (*Amphilophus* spp.) in the Nicaraguan crater lakes. *BMC Evol. Biol.* 10: 326.

Berner, D., Grandchamp, A.-C. & Hendry, A.P. 2009. Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. *Evolution* 63: 1740–53.

Betancur-R., R. 2009. Molecular phylogenetics and evolutionary history of ariid catfishes revisited: a comprehensive sampling. *BMC Evol. Biol.* 9: 175.

Betancur-R., R. 2010. Molecular phylogenetics supports multiple evolutionary transitions from marine to freshwater habitats in ariid catfishes. *Mol. Phylogenet. Evol.* 55: 249–258. Elsevier Inc.

Betancur-R., R. & Acero, A. 2004. Description of *Notarius biffi* n. sp. and redescription of *N. insculptus* (Jordan and Gilbert) (Siluriformes: Ariidae) from the eastern Pacific, with evidence of monophyly and limits of *Notarius*. *Zootaxa* 703: 1–20.

Betancur-R., R. & Acero P., A. 2005. Description of *Cathorops mapale*, a new species of sea catfish (Siluriformes: Ariidae) from the Colombian Caribbean, based on morphological and mitochondrial evidence. *Zootaxa* 1045: 45–60.

Betancur-R., R., Acero P., A., Bermingham, E. & Cooke, R. 2007. Systematics and biogeography of New World sea catfishes (Siluriformes: Ariidae) as inferred from mitochondrial, nuclear, and morphological evidence. *Mol. Phylogenet. Evol.* 45: 339–357.

Betancur-R., R., Marceniuk, A.P. & Béarez, P. 2008. Taxonomic Status and Redescription of the Gillbacker Sea Catfish (Siluriformes: Ariidae: *Sciades parkeri*). *Copeia* 4: 827–834.

Betancur-R., R., Orti, G., Stein, A.M., Marceniuk, A.P. & Pyron, R.A. 2012. Apparent signal of competition limiting diversification after ecological transitions from marine to freshwater habitats. *Ecol. Lett.* 15: 822–830.

Bloch, J.I., Woodruff, E.D., Wood, A.R., Rincon, A.F., Harrington, A.R., Morgan, G.S., *et al.* 2016. First North American fossil monkey and early Miocene tropical biotic interchange. *Nature* 533: 243–246.

Bloom, D.D., Weir, J.T., Piller, K.R. & Lovejoy, N.R. 2013. Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analyses and molecular phylogenetics of New World silversides (Atherinopsidae). *Evolution.* 67: 2040–2057.

Carrillo, J.D., Forasiepi, A., Jaramillo, C. & Sánchez-Villagra, M.R. 2015. Neotropical mammal diversity and the Great American Biotic Interchange: spatial and temporal variation in South Amercia's fossil record. *Front. Genet.* 5: 451.

da Silva, W.C., Marceniuk, A.P., Sales, J.B.L. & Araripe, J. 2016. Early Pleistocene lineages of *Bagre bagre* (Linnaeus, 1766) (Siluriformes: Ariidae), from the Atlantic coast of South America, with insights into the demography and biogeography of the species. *Neotrop. Ichthyol.* 14: e150184.

Davis, A.M., Unmack, P.J., Pusey, B.J., Johnson, J.B. & Pearson, R.G. 2012. Marine-freshwater transitions are associated with the evolution of dietary diversification in terapontid grunters (Teleostei: Terapontidae). *J. Evol. Biol.* 25: 1163–1179.

de Queiroz, A. 2014. *The Monkey's Voyage: How Improbable Journeys Shaped the History of Life*. Basic Books, Perseus Books Group, New York.

Farris, D.W., Jaramillo, C., Bayona, G., Restrepo-Moreno, S.A., Montes, C., Cardona, A., *et al.* 2011. Fracturing of the Panamanian Isthmus during initial collision with: South America. *Geology* 39: 1007–1010.

Froese, R. & Pauly, D. 2016. FishBase. World Wide Web electronic publication. www.fishbase.org, (10/2016)

Hohenlohe, P., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E. & Cresko, W. A. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* 6: e1000862.

Kimmel, C.B., Cresko, W.A., Phillips, P.C., Ullmann, B., Currey, M., von Hippel, F., *et al.* 2012. Independent axes of genetic variation and parallel evolutionary divergence of opercle bone shape in threespine stickleback. *Evolution.* 66: 419–434.

Knowlton, N., Weigt, L.A., Solorzano, L.A., Mills, D.K. & Bermingham, E. 1993. Divergence in Proteins, Mitochondrial-DNA, and Reproductive Compatibility across the Isthmus of Panama. *Science.* 260: 1629–1632.

Lessios, H.A. 1981. Divergence in Allopatry : Molecular and Morphological Differentiation Between Sea Urchins Separated by the Isthmus of Panama. *Evolution*. 35: 618–634.

Lovejoy, N.R. & Collette, B.B. 2001. Phylogenetic relationships of New World needlefishes (Teleostei: Belonidae) and the biogeography of transitions between marine and freshwater habitats. *Copeia* 324–338.

Marceniuk, A.P. 2007. Revalidação de *Cathorops arenatus* e *Cathorops agassizii* (Siluriformes, Ariidae), bagres marinhos das regiões norte e nordeste da América do Sul. *Iheringia, Série Zool.* 97: 360–375.

Marceniuk, A.P., Betancur-R., R., Acero P., A. & Muriel-Cunha, J. 2012. Review of the Genus *Cathorops* (Siluriformes: Ariidae) from the Caribbean and Atlantic South America, with Description of a New Species. *Copeia* 2012: 77–97.

Marceniuk, A.P. & Menezes, N. 2007. Systematics of the family Ariidae (Ostariophysi, Siluriformes), with a redefinition of the genera. *Zootaxa* 1416: 1–126.

Marceniuk, A.P., Menezes, N.A. & Britto, M.R. 2012. Phylogenetic analysis of the family Ariidae (Ostariophysi: Siluriformes), with a hypothesis on the monophyly and relationships of the genera. *Zool. J. Linn. Soc.* 165: 534–669.

Martin, A.P., Naylor, G.J.P. & Palumbi, S.R. 1992. Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. *Nature* 357: 153–155.

Matschiner, M., Hanel, R. & Salzburger, W. 2011. On the origin and trigger of the notothenioid adaptive radiation. *PLoS One* 6: e18911.

Montes, C., Bayona, G., Cardona, A., Buchs, D.M., Silva, C.A., Morón, S., *et al.* 2012. Arc-continent collision and orocline formation: Closing of the Central American seaway. *J. Geophys. Res. Solid Earth* 117: 1–25.

Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J.C., Valencia, V., *et al.* 2015. Middle Miocene closure of the Central American Seaway. *Science (80-.).* 348: 226–229.

O'Dea, A., Lessios, H.A., Coates, A.G., Eytan, R.I., Restrepo-Moreno, S.A., Cione, A.L., *et al.* 2016. Formation of the Isthmus of Panama. *Sci. Adv.* 2: e1600883.

Rabosky, D.L., Santini, F., Eastman, J., Smith, S. a, Sidlauskas, B., Chang, J., *et al.* 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat. Commun.* 4: 1958.

Rutschmann, S., Matschiner, M., Damerau, M., Muschick, M., Lehmann, M.F., Hanel, R., *et al.* 2011. Parallel ecological diversification in Antarctic notothenioid fishes as evidence for adaptive radiation. *Mol. Ecol.* 20: 4707–4721.

Seehausen, O., Terai, Y., Magalhaes, I.S., Carleton, K.L., Mrosso, H.D.J., Miyagi, R., *et al.* 2008. Speciation through sensory drive in cichlid fish. *Nature* 455: 620–626.

Stange, M., Aguirre-Fernandez, G., Cooke, R.G., Barros, T., Salzburger, W. & Sánchez-Villagra, M.R. 2016. Evolution of opercle bone shape along a macrohabitat gradient: species identification using mtDNA and geometric morphometric analyses in neotropical sea catfishes (Ariidae). *Ecol. Evol.* 6: 5817–5830.

Teugels, G.G. 1996. *Taxonomy, phylogeny and biogeography of catfishes (Ostariophysi, Siluroidei): an overview* (M. Legendre & J.-P. Proteau, eds). Aquatic Living Resources.

Wilson, L.A.B., Colombo, M., Hanel, R., Salzburger, W. & Sánchez-Villagra, M.R. 2013. Ecomorphological disparity in an adaptive radiation: opercular bone shape and stable isotopes in Antarctic icefishes. *Ecol. Evol.* 3: 3166–3182.

Wilson, L.A.B., Colombo, M., Sánchez-Villagra, M.R. & Salzburger, W. 2015. Evolution of opercle shape in cichlid fishes from Lake Tanganyika - adaptive trait interactions in extant and extinct species flocks. *Sci. Rep.* 5: 16909. Nature Publishing Group.

Woodburne, M.O. 2010. The Great American Biotic Interchange: Dispersals, Tectonics, Climate, Sea Level and Holding Pens. *J. Mamm. Evol.* 17: 245–264.

Chapter 1

Evolution of opercle bone shape along a macrohabitat gradient: species identification using mtDNA and geometric morphometric analyses in Neotropical sea catfishes (Ariidae)

Ecology and Evolution

Open Access

Evolution of opercle bone shape along a macrohabitat gradient: species identification using mtDNA and geometric morphometric analyses in neotropical sea catfishes (Ariidae)

Madlen Stange[1,2], Gabriel Aguirre-Fernández[1], Richard G. Cooke[3], Tito Barros[4], Walter Salzburger[2] & Marcelo R. Sánchez-Villagra[1]

[1]Palaeontological Institute and Museum, University of Zurich, Karl-Schmid-Strasse 4, 8006, Zurich, Switzerland
[2]Zoological Institute, University of Basel, Vesalgasse 1, 4051, Basel, Switzerland
[3]Smithsonian Tropical Research Institute, MRC 0580-08, Apartado, 0843-03092, Panama, Republic of Panama
[4]Museo de Biología, Facultad Experimental de Ciencias, La Universidad del Zulia, Apartado Postal 526, Maracaibo, 4011, Estado Zulia, Venezuela

Keywords
Geometric morphometrics, macrohabitat transition, mitochondrial DNA, Siluriformes, systematics, taxonomy.

Correspondence
Madlen Stange and Marcelo R. Sánchez-Villagra, Palaeontological Institute and Museum, University of Zurich, Karl-Schmid-Strasse 4, 8006 Zurich, Switzerland.
Tel: +41 (0)44 634 23 38;
Fax +41 (0)44 634 49 23;
E-mail: madlen.stange@pim.uzh.ch (M.S.)
Tel: +41 (0)44 634 23 42;
Fax: +41 (0)44 634 49 23;
E-mail: m.sanchez@pim.uzh.ch (M.R.S.-V.)

Funding Information
Forschungskredit of the University of Zurich, (Grant/Award Number: "FK-15-092")
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, (Grant/Award Number: "CRSII3-136293").

Received: 10 June 2016; Accepted: 29 June 2016

Ecology and Evolution **2016; 6(16): 5817–5830**

doi: 10.1002/ece3.2334

Abstract

Transitions between the marine and freshwater macrohabitat have occurred repeatedly in the evolution of teleost fishes. For example, ariid catfishes have moved from freshwater to marine environments, and vice versa. Opercles, a skeletal feature that has been shown to change during such transitions, were subjected to 2D geometric morphometric analyses in order to investigate evolutionary shape changes during habitat transition in ariid catfishes and to test the influence of habitat on shape changes. A mtDNA marker, which proved useful in previous studies, was used to verify species identities. It greatly improved the assignment of specimens to a species, which are difficult to assign by morphology alone. The application of a mtDNA marker confirmed the occurrence of *Notarius biffi* in Central America, South of El Salvador. Molecular identification together with principal component analysis (PCA) and further morphological inspection of neurocrania indicated the existence of a cryptic species within *Bagre pinnimaculatus*. Principal component (PC) scores of individual specimens clustered in morphospace by genus rather than by habitat. Strong phylogenetic structure was detected using a permutation test of PC scores of species means on a phylogenetic tree. Calculation of Pagel's λ suggested that opercle shape evolved according to a Brownian model of evolution. Yet canonical variate analysis (CVA) conducted on the habitat groups showed significant differences in opercle shapes among freshwater and marine species. Overall, opercle shape in tropical American Ariidae appears to be phylogenetically constrained. This verifies the application of opercle shape as a taxonomic tool for species identification in fossil ariid catfishes. At the same time, adaptation to freshwater habitats shows characteristic opercle shape trajectories in ariid catfishes, which might be used to detect habitat preferences in fossils.

Introduction

The colonization of freshwater habitats by marine ancestors (and vice versa) is a common scenario in the evolution of fishes and has occurred multiple times in various teleost families, including Gasterosteidae (sticklebacks) (Bell and Foster 1994), Belonidae (needlefish) (Lovejoy et al. 2006), Atherinopsidae (neotropical silversides) (Bloom et al. 2013), and Terapontidae (grunters) (Davis et al. 2012), as well as in elasmobranchs, such as Myliobatiformes (stingrays) (Lovejoy et al. 2006) and Carcharhinidae (requiem sharks) (de Carvalho, M.R., McEachran, J.D., 2003). Such transitions from marine to freshwater macrohabitats implicate varied adaptations to different environmental conditions in, for example, salinity, pH, flow conditions, oxygen content, and feeding

ecology. Here, we investigated the opercle, a skeletal feature of teleosts that is not actively involved in the adaptation to any of these changes but may be indirectly affected, and hence should reflect adaptations to a variety of environmental changes. The opercle is positioned laterally and centrally on the head, and changes dimensions according to skull shape. Its primary function is the protection of the gills, which make it useful for study as it is present in most teleost fishes, it can be easily examined from the outside, and it is often well preserved in the fossil record. Therefore, the study of the opercle allows comparisons both across teleosts in general and in deep time (see, e.g., Wilson et al. 2015).

Opercle shape has been studied in different contexts in various fish taxa, including extant Alaskan *Gasterosteus aculeatus* (stickleback) populations, Antarctic notothenioids, Lake Tanganyika cichlids, and extinct *Saurichthys* species (Kimmel et al. 2008; Wilson et al. 2013a,b, 2015); the development of the opercle has further been studied in *Danio rerio* (zebrafish) (Kimmel et al. 2010) and in *Saurichthys* (Scheyer et al. 2014). Overall, these studies have revealed a strong link between the shape of the opercle and the habitat and/or feeding type of the respective species. For example, Alaskan sticklebacks that invaded lake environments diverged from their anadromous ancestors in opercle shape (Kimmel et al. 2005, 2012a; Arif et al. 2009) with the opercle of lake populations being stretched along the anterior–posterior axis, while compressed along the dorsal–ventral axis (Kimmel et al. 2008). The characteristic shape of the opercle in freshwater specimens is likely due to a developmental decoupling of shape and size, as freshwater populations retain the juvenile ancestral opercle shape despite achieving full body size (Kimmel et al. 2012b). Indeed, opercle size and shape have previously been demonstrated to arise from two independent developmental modules in *D. rerio* (Kimmel et al. 2010). Divergence in opercle shape has further been shown between inhabitants of shallow and deep Alaskan lakes (Arif et al. 2009). Among extinct marine *Saurichthys* species, a dorsoventral compression could be observed (Wilson et al. 2013b). Similar to *G. aculeatus*, but slightly offset, an anterior–posterior stretching and dorsal–ventral compression was observed in Antarctic notothenioid fishes (Wilson et al. 2013a). The notothenioids, which diverged between 17.1 and 10 mya (Colombo et al. 2015), represent a relatively old adaptive radiation, especially when compared to sticklebacks. Another pattern observed in notothenioids is that opercle shape is strongly influenced by phylogeny, although within subfamilies, opercle shapes reflect adaptations along the benthic–pelagic axis (Wilson et al. 2013a). In the cichlid adaptive radiation in Lake Tanganyika, a significant correlation between opercle shape and gut length was uncovered (Wilson et al. 2015), suggesting that opercle shape reflects adaptations in feeding ecology. In summary, two major factors have been identified to contribute to opercle shape: habitat and phylogeny (Wilson et al. 2015).

Here, we investigated opercle shape in neotropical sea catfishes of the "family" Ariidae and the influence of habitat and phylogeny upon it. Ariid catfishes belong to an exceptional fish radiation (Alfaro et al. 2009) that diversified along a marine–freshwater habitat trajectory. These fish may therefore show evidence of opercle shape evolution along this macrohabitat axis, comprising adaptation events in recent as well as in geological times. Siluriformes (catfishes) diverged sometime between 180 mya (Nakatani et al. 2011 [molecular data]) and before the Late Cretaceous (more than 100 mya) (Diogo 2004 [fossil evidence]). Today the Siluriformes consist of more than 3000 species [www.fishbase.org, version 01/2016]) in 33 "families", with most of the catfish species being primary freshwater inhabitants (Teugels 1996). Some "families" contain species with a preference for brackish habitats, such as the Loricariidae and Pimelodidae (Betancur-R 2009, 2010), but only two of the 33 "families" – including the widely distributed Ariidae – can be characterized as primarily marine. Nonetheless, some members of the Ariidae have secondarily colonized freshwater habitats, so that Ariidae inhabit the coastal waters and near-coastal rivers and lakes of most tropical and subtropical regions worldwide (Sullivan et al. 2006). The evolutionary history of Ariidae has been relatively well traced in tropical America, as some structures such as the lapillus (largest ear stones in Ariidae), spines, and partial skulls preserve rather well in the fossil record. All these features contain taxonomic information (Aguilera et al. 2013), with the lapillus in particular allowing identification to the genus and in some cases species level (Acero and Betancur-R 2007 and references therein; Chen et al. 2011).

A habitat change from marine to freshwater, as recorded in the Ariidae, was a fundamental part of the history of other South American biota, including dolphins, stingrays, and needlefish (Lovejoy et al. 2006). During a major event in the Miocene (23–5 mya), sea level differences and tectonic activities shaped and reshaped the aquatic landscape on this continent, forming the Amazon River (Hoorn and Hoorn 2006; Lovejoy et al. 2006). Some tropical American ariid species are known to be present in their respective habitats since the lower Miocene (23 mya) (Aguilera and de Aguilera 2004). Other ariid species might have taken the opportunity to occupy new habitats during these major geological events. Indeed, it is likely that the Ariidae still have the propensity to occupy freshwater habitats as it has been observed in *Cathorops tuyra* in Panama (D. Sharpe, pers. comm.,

December, 2015). During our sampling, we found species in habitats and localities in which they had not been reported to occur. For that reason, we summarized the species we found in each habitat and substantiated the correctness of species identity by the usage of a mtDNA marker for comparison with a reference dataset.

We investigated habitat transition in a group of catfishes by providing information on distribution and taxonomy as examined by molecular markers in a phylogenetic context and based on shape analysis of the opercle bone as morphological marker. Unlike Antarctic notothenioids, Ariidae show no subgroup divergence within genera. Nonetheless, opercle shapes between marine and freshwater groups are significantly different. The same pattern of dorsal–ventral compression and anterior–posterior stretching along the two major axes of variance was detected as in the *Gasterosteus aculeatus* (stickleback), Antarctic notothenioids, and Lake Tanganyika cichlids, suggesting this to be a general trend in teleosts.

Methods

Sampling and species Identification using the mtDNA marker *ATPase 8/6*

Twenty-two ariid catfish species belonging to the genera *Bagre*, *Sciades*, *Cathorops*, *Notarius*, and *Ariopsis* were sampled in marine, brackish, and freshwater habitats of Venezuela (VE) and Panama (PA) (Fig. 1). GPS coordinates, sampled species in each location, and habitat definition can be found in Table S1. In Venezuela, specimens were bought fresh from local fish markets, or from hired fishers, while in Panama specimens were caught directly by the authors with the support of local fishers. Taxonomic identifications of all specimens were conducted in the field on the basis of characteristics of morphological traits, for example, neurocranium shape (visible through the skin), color, dentary morphology and, after maceration, neurocrania again (Fischer et al. 1995; Acero 2002;

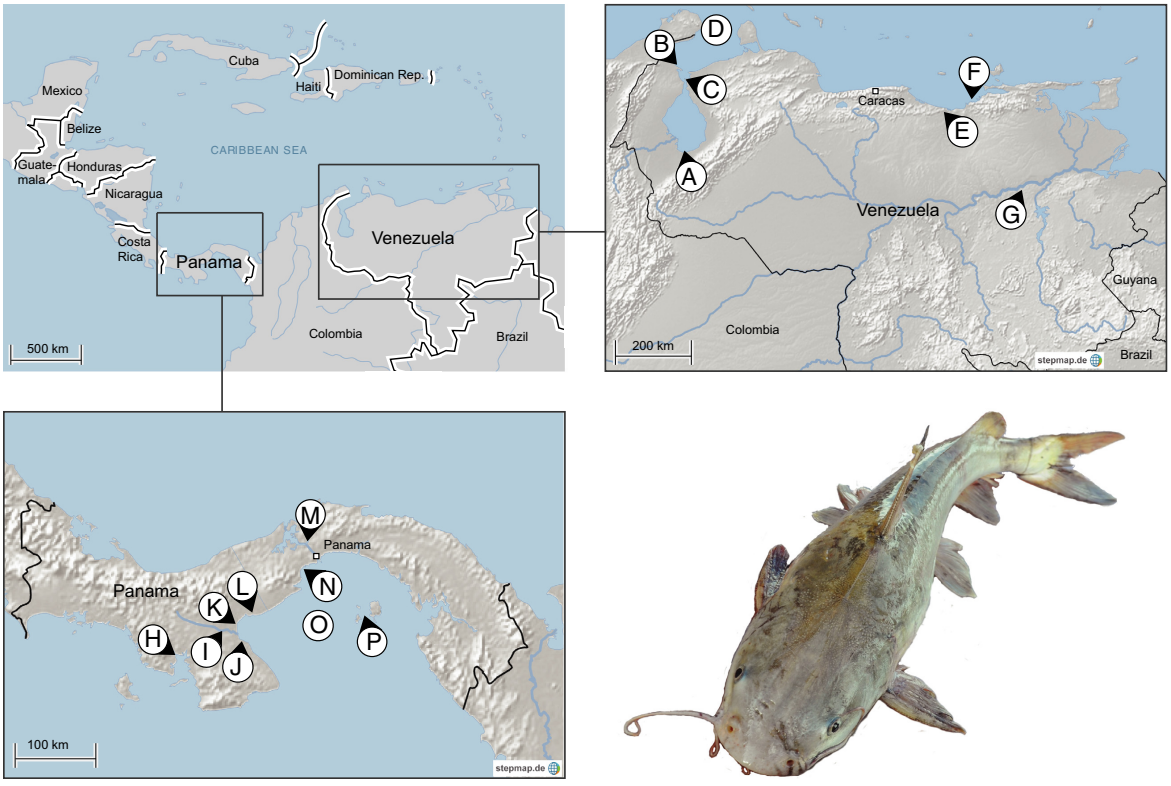


Figure 1. Map of sampling locations. The fish picture depicts a representative of *Cathorops*. A: Puerto Concha, Maracaibo Lake, Zulia state, VE. B: Isla de Toas and Isla de San Carlos, Maracaibo Lake, Zulia state, VE. C: Guarico, Maracaibo Lake, Zulia state, VE. D: Gulf of Venezuela, Falcón state, VE. E: Clarines, Anzoátegui state, VE. F: Puerto la Cruz, Anzoátegui state, VE. G: Ciudad Bolivar, Bolivar state, VE. H: Río San Pedro, Montijo Bay, PA. I: Rio Santa Maria, PA. J: Rio Parita, PA. K: Rio Estero Salado, PA. L: Rio Hato, PA. M: Rio Chagres, PA. N: Puerto Caimito, PA. O: Gulf of Panama, PA. P: Pearl Islands, Casaya Island, PA. For additional information, see Tables S1 and S3.

Betancur-R and Acero 2005; Marceniuk 2007; Lasso and Sánchez-Duarte 2011; Marceniuk et al. 2012). The sampled material included caudal fin tissue preserved in 96% ethanol for DNA isolation, and macerated and bleached opercle bones. For verification of the taxonomic affiliation, sequencing of a mtDNA marker was conducted.

In total, 263 ariid catfishes were sampled, and isolated DNA (for unique sample identifiers with species, locality information and NCBI accession numbers, see Table S3) was subjected to standard Sanger sequencing targeting mitochondrial *ATPase 8/6*. DNA was isolated using standard salt precipitation. For details on amplification and sequencing, see Appendix S1. Sequence traces (available on NCBI, accession numbers KX500399-KX500661) were inspected for quality control, removing poor-quality bases and/or ambiguous base calls using CodonCode Aligner v.3.7.1.1 (CodonCode, Dedham, MA). These were aligned using MAFFT v.7 (Katoh and Standley 2013) together with a reference dataset (Betancur-R 2009) containing 281 *ATPase 8/6* sequences from 129 ariid species distributed worldwide. This reference dataset was kindly provided by R. Betancur-R. Maximum-likelihood trees were constructed using RAxML 7.0.3 (Stamatakis 2006) under the GTR+I+G model. The best tree from 1000 bootstrap replicates was used to infer molecular species identity by clustering of our sampled specimens to the reference sequences of species.

2D geometric morphometric (GM) data collection and analyses

Opercle bones of 263 specimens (deposited at the Palaeontological Institute and Museum of the University of Zurich, PIMUZ; for unique identifiers, see Table S3) were subjected to GM analysis. Left and right opercles (if present) of each specimen were photographed from lateral view using a digital camera mounted on a table stand. Raw images of the right opercle were reorientated and reflected in Adobe Photoshop CS6 to match left opercle orientation. This is necessary for merging the shape data of left and right opercle in downstream analyses. All raw images were further processed using the TpsUtil v.1.60 software (http://life.bio.sunysb.edu/morph/soft-utility.html). In order to analyze the shape of the opercle bone, a sliding semilandmark approach was applied. The first landmark was defined as the most ventral point of the anterior edge and represents the only true or fixed landmark (Type 2 landmark) in this study. Ninety-nine equidistant points (Fig. 2) were then placed, capturing the outline of each opercle using TpsDig v2.10 (Rohlf 2013). This approach was chosen above a true landmark approach, as the opercle shapes among genera varied greatly and prohibited the identification of further homologous landmarks. During Procrustes superimposition, performed in the R

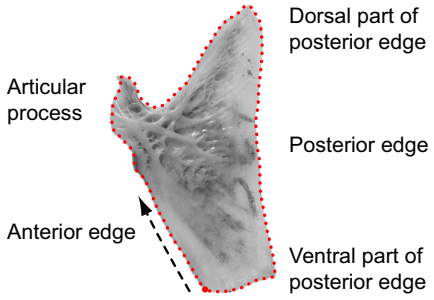


Figure 2. One hundred equidistant landmarks on a left ariid opercle (*Notarius quadriscutis*). The starting and end point of landmark capturing is indicated by a bold red dot, representing a Type 2 landmark, the local maximum of the curvature on the ventral part of the anterior edge. Ninety-nine sliding semilandmarks were equally spaced capturing the outline of the opercle bone. Direction of landmark recording along the outline is indicated by an arrow.

package geomorph v.3.0.0 (Adams and Otarola-Castillo 2013), differences in scale, orientation, and size were removed, and each semilandmark was moved along its tangent vector (between its left and right neighboring semilandmark) until its position minimized the shape difference between specimens based on bending energy (Bookstein 1997; Bookstein et al. 1999; Gunz et al. 2005; Rohlf 2010). The slid and superimposed landmark coordinates were imported to MorphoJ 1.06d (Klingenberg 2011), where all subsequent morphometric analyses were conducted.

The shape differences between left and right opercle bones within specimens were assessed using Procrustes ANOVA on geometric shape data (Klingenberg et al. 2002). Interspecimen shape differences were much larger than intraspecimen shape differences (within-specimen Procrustes sum of squares = 0.0021 [F = 1.9, P = <0.0001]; between-specimen Procrustes sum of squares = 2.88 [F = 28.24, P = <0.0001]). Therefore, we averaged shape data (Procrustes coordinates and centroid size) of left and right opercle by specimen for downstream analyses if both the left and right opercles were available (n = 250), or included shape data for single opercles if only one opercle was available (n = 13).

Allometric effects within species, caused by specimens of different ontogenetic stages, were removed by regressing pooled within-species Procrustes coordinates (shape) onto centroid size (Drake and Klingenberg 2008). The resulting regression residuals were then used to calculate a covariance matrix for the subsequent multivariate analyses.

Patterns of opercle shape variation among ariid species

Principal component analyses (PCAs) were used to identify axes of maximal shape variance among all ariid

specimens in order to discover patterns of variation, to explore groupings among them, and to assess phylogenetic structure in the groupings.

In a first PCA (Fig. 3), principal component (PC) scores for all 263 samples were calculated and plotted to investigate the clustering behavior of our samples in morphospace. The number of individuals per species and locality is listed in Table S2. Subsequently, we tested for phylogenetic structure in the shape data by calculating centroid sizes and Procrustes coordinates, representing species means, in order to match shape and size values with the terminal taxa in the phylogenetic tree. A new covariance matrix was calculated, a second PCA was performed, and a time-calibrated tree constructed from mitochondrial and nuclear markers, published by Betancur-R et al. (2012), was mapped onto the PC scores (Fig. 4A). This time-calibrated tree (Betancur-R et al. 2012) represents the currently most resolved phylogenetic tree of the Ariidae. Although the tree contains a large number of species, not all species analyzed here are contained, forcing us to average shape data of *C. nuchalis* and *C. wayuu* to *Cathorops* sp. This restriction also hindered our analysis of within-species shape differentiation (among habitats) as data for freshwater and marine

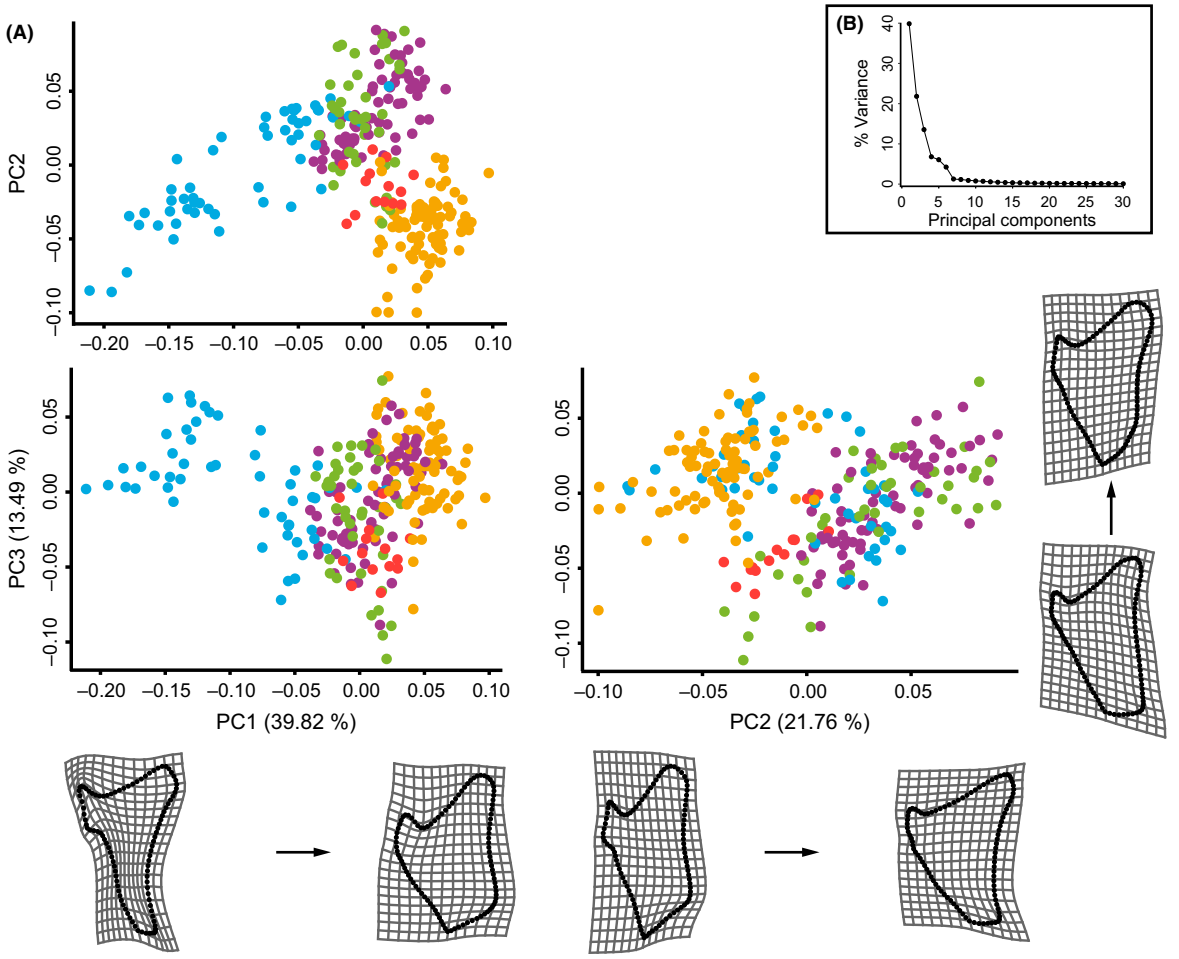


Figure 3. Patterns of ariid opercle shape variation along major axes of variance. (A) Scatterplots of the first three principal components (PCs) of ariid opercle shapes (N = 263) from 21 species belonging to the genera *Bagre* (●), *Sciades* (●), *Cathorops* (●), *Notarius* (●), and *Ariopsis* (●). A detailed scatterplot of species can be found in Figure S1. The opercle shape changes are displayed below or next to the respective axes of variance using thin plate spline visualization grids using the starting shape at scale factor 0 as reference shape. The respective shape changes are scaled by −0.2 and +0.1 (magnitude of shape change in Procrustes distance) for PC1, −0.1 and +0.1 for PC2, and −0.1 and +0.07 for PC3. (B) Scree plot of the first 30 PCs and their amount of variance. The bent after the third PC indicates a distinct drop in the impact of following PCs on shape variance. Therefore, only the first three PCs are presented in the scatterplot. The first three PCs together account for 75.07% of the observed variation.

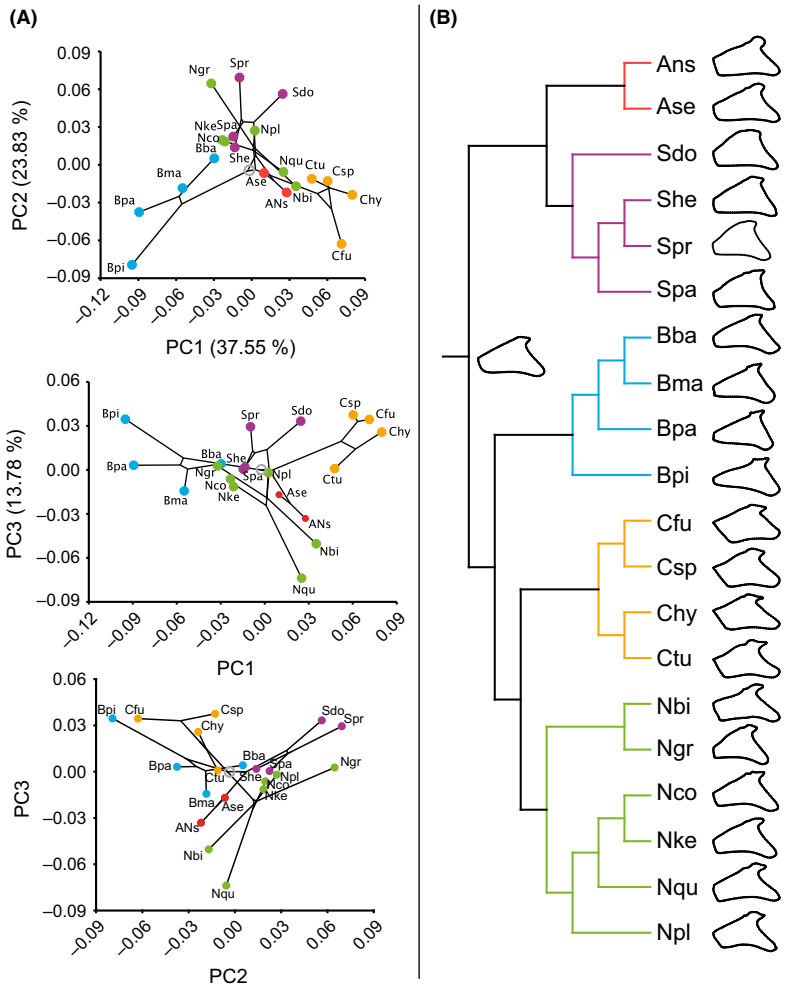


Figure 4. Evolutionary opercle shape change within Ariidae. (A) Phylomorphospace plot of ariid opercle shape changes. A time-calibrated tree constructed from mitochondrial and nuclear markers published by Betancur-R et al. (2012) has been projected on PC scores derived from species means. The first three principal components account for 75.16% of the total variation. PC scores of 20 species are displayed and highlighted by genus (*Bagre* (●), *Sciades* (●), *Cathorops* (●), *Notarius* (●), and *Ariopsis* (●), root (○)). Ans: *A.* sp. nov. Ase: *A. seemanni*; Sdo: *S. dowii*; She: *S. herzbergii*; Spr: *S. proops*; Spa: *S. parkeri*; Bba: *B. bagre*; Bma: *B. marinus*; Bpa: *B. panamensis*; Bpi: *B. pinnimaculatus*; Cfu: *C. fuerthii*; Csp: *Cathorops* sp. (includes *C. wayuu* and *C. nuchalis*, as both species do not differ in the 11 genes analyzed by Betancur-R et al.); Chy: *C. hypophthalmus*; Ctu: *C. tuyra*; Nbi: *N. biffi*; Ngr: *N. grandicassis*; Nco: *N. cookei*; Nke: *N. kessleri*; Nqu: *N. quadriscutis*; Npl: *N. planiceps*. PC scores for *C.* sp. indet. are not displayed, as they are not included in the phylogeny of Betancur-R et al. (2012). The permutation of shapes along the phylogeny resulted in a *P*-value of 0.001 rejecting the null hypothesis of absence of phylogenetic signal. Pagel's λ of 1 is not significantly different from 1 ($P = 1$) implying opercle shape evolution happened according to the Brownian model of evolution. (B) Evolutionary opercle shape change along the time-calibrated tree (Betancur-R et al. 2012) that has been projected on Procrustes coordinates derived from species means. Mean shapes per species are displayed at the node tips, and the ancestral opercle state for the shapes studied here is displayed at the first internal split at the root.

populations had to be merged for phylogenetically corrected shape analyses. A total of 1000 permutations simulating the null hypothesis of total absence of phylogenetic structure by randomly permuting the PC scores among the species and mapping onto the phylogeny were applied (Klingenberg and Gidaszewski 2010). The resulting *P*-value represents the proportion of permutations that resulted in equal or shorter tree lengths than the input tree. To additionally test for phylogenetic dependence on the species trait values, in this case PC scores per species, a phylogenetic least squares (PGLS) analysis (Martins and Hansen 1997) was performed using a maximum-likelihood estimate of Pagel's λ, implemented in the R package "Caper" (Orme et al. 2013). This test assumes a Brownian

model of evolution, where variation between tips along all branches of the tree accumulates at a rate proportional to the length of the branches, meaning that the more closely that taxa are related, the more similar they are in trait values, and vice versa ($\lambda = 1$). A trait value evolution that disagrees with the Brownian model would result in $\lambda = 0$. The evolutionary opercle shape change and ancestral shape were reconstructed by projecting the phylogeny on Procrustes coordinates of species means (Fig. 4B).

Opercle shape variation according to habitat

Canonical variate analysis (CVA) was used to visualize opercle shape changes that discriminate among the different habitats of our sampled specimens. CVA computes axes of variance in a way that minimized within-group differences and maximized between-group differences. Samples were assigned a priori to the following groups: freshwater ($n = 38$, 5 species), brackish ($n = 94$, 14 species), or marine ($n = 131$, 8 species). For numbers of individuals used per species and habitat, see Table S2. As the precise habitat range of most species is unknown, the classification as freshwater-, brackish-, or marine-occurring species follows the habitat where fishers or the authors captured the specimens. Therefore, individuals of the same species were in some cases assigned to different habitats. As in the PCA, the residuals of the within-group regression of Procrustes coordinates onto centroid size per specimen were used (in order to correct data for allometry). The significance of differences among group means (habitats) was tested in a permutation test with 1000 random permutations, and distances are presented in Procrustes and Mahalanobis measures.

Results

Species identification using the mtDNA marker *ATPase 8/6*

The morphological and molecular assignment of specimens was not concurrent in all cases. By aligning gained *ATPase 8/6* sequences against a reference dataset, 19 ariid species instead of 21, from five genera, were confirmed. Four individuals had identical sequences but could not be assigned to any specific species, using either morphological or molecular data. These sequences showed close affinity to *Cathorops*. The individuals were therefore labeled as *C.* sp. indet. Some species that originated from different localities exhibited distinct *ATPase 8/6* sequences; others had identical sequences despite different sampling localities, and still others were found in unreported habitats. The results are summarized in Table 1 (see also Appendix S2).

Geometric morphometrics

Patterns of opercle shape variation among ariid species

Principal component analysis (PCA) was used to investigate the distribution of opercle shapes of ariid catfishes from different habitats in morphospace. Despite being small (3.8%), the within-group size variation was removed prior to PCA.

The first three PCs accounted for 75.07% of the observed variation (Fig. 3B). In the morphospace plot, *Bagre* took negative values along PC1 (Fig. 3A) with *B. pinnimaculatus* separating in morphospace (most left cluster) from the other *Bagre* species (Fig. S1), as well as from all other analyzed species. The two *B. pinnimaculatus* populations (inferred from typical hyperossified frontals vs. lack of hyperossified frontals, as well as different *ATPase 8/6* haplotypes) did not overlap in the scatterplot of PC1 versus PC2 and PC1 versus PC3 (data not presented here), with *B. pinnimaculatus* featuring the hyperossification of the frontals forming the lower part of the *B. pinnimaculatus* cohort. PC2 did not distinguish any of the groups in the scatterplot. The least overlap of genera and species in opercle shape was evident in the scatterplot of PC1 versus PC2 (Fig. 3A). *Bagre* and *Cathorops* were separated along PC1, whereas the genera *Ariopsis*, *Sciades*, and *Notarius* overlap in morphospace. The opercle shape changes along PC1 and PC2 followed a dorsal–ventral compression and an anterior–posterior stretching, respectively (Fig. 3A). PC3 demonstrated the shape change from a blunt ventral surface of the posterior edge to a pointy ventral surface.

It became apparent that opercle shapes of the specimens cluster by genus rather than by habitat. To test for phylogenetic signal in our shape data, a phylogeny was plotted on PC scores of species means producing a phylomorphospace. Genera took mainly distinct areas in the phylomorphospace, deviating from the root (Fig. 4A). An apparent phylogenetic structure was tested with the null hypothesis of absence of phylogenetic signal in the morphometric shape data and was rejected ($P = 0.001$). Additionally, Pagel's λ of 1 tested on PC1 is not significantly different from 1 ($P = 1$); likewise, λ for PC2 is 0.75 ($P = 0.19$), suggesting that opercle shape evolution happened according to the Brownian model and that opercle shape traits are as similar as expected by their phylogenetic distance. The result that opercle shapes were most similar within each genus is visualized in the plot of Procrustes coordinates (shape) on the phylogeny and the ancestral shape reconstruction (Fig. 4B).

Table 1. Species identification using a mtDNA marker *ATPase 8/6*.

Genus	Species	Localities[1]	Notes
Sciades	*proops* (Valenciennes, 1840)	D^m (17), E^m (1), F^m (1)	Two sequence clusters corresponding to sampling locations, E and F, differ from D
	herzbergii (Bloch, 1784)	D^m (31), E^m (15)	Two sequence clusters corresponding to sampling locations
	dowii (Gill 1863)	Mouth of L^b (1), N^m (2), I^f (2)	No sequence difference according to sampling location
	parkeri (Traill 1832)	G^f (2)	This species is described as occurring in lower parts of rivers from the Gulf of Paria, VE, to Brazil (Betancur-R et al. 2008). In this study, the species was found 320 km inland
Ariopsis	*seemanni* (Günther 1864)	K^b (5)	
	sp. nov. (*A. jimenezi*, A. Marceniuk et al., in prep.)	P^m (9)	
Notarius	*cookei* (Acero and Betancur-R 2002)	I^f (7)	Described as brackish water species (Betancur-R et al. 2007). In this study, it was sampled in freshwater, 10 km upstream from the mouth of the Santa Marias river. Vega (pers. comm., Dec. 2015) captured it in the same river at locations between 69 and 76 lm from the mouth in completely fresh water
	kessleri (Steindachner 1876)	H^b (8), K^b (2)	No sequence difference according to sampling location
	biffi (Betancur-R and Acero 2004)	H^b (1)	First reported as restricted from El Salvador to Costa Rica (Betancur-R and Acero 2004), but see Robertson, 2015 who presented records from western and central Panama to Parita Bay. We sampled it in the Gulf of Montijo, Panama
	quadriscutis (Valenciennes, 1840)	E^b (10)	
	grandicassis (Valenciennes, 1840)	D^m (11)	
	planiceps (Steindachner, 1877)	K^b (1)	
Bagre	*bagre* (Linnaeus, 1766)	D^m (3), B^b (2)	No sequence difference according to sampling location.
	aff. *marinus*	D^m (16), F^m (3)	No sequence difference according to sampling location
	panamensis (Gill 1863)	K^b (3)	
	pinnimaculatus (Steindachner 1876)	K^b (1), O^m (4, atypical: 20)	Several specimens (atypical) from the Gulf of Panama (O) diverged from the morphology of *B. pinnimaculatus* by lacking the typical hyperossification of the frontals, preopercle, and interopercle. They did not exhibit the phenotype of *B. panamensis*, either, the only other *Bagre* species occurring in the eastern Pacific. Morphologically they resembled *B. bagre* from the Atlantic. In the ML tree, those specimens formed a sister clade to *B. pinnimaculatus*. However, due to their molecular relatedness they were considered being *B. pinnimaculatus* in the subsequent analyses. Examination of additionally sampled eleven specimens at Puerto Coquira, Panama Bay, PA (not included in this study), showed that only three individuals had hyperossified frontals and eight were lacking the typical hyperossification
Cathorops	*hypophthalmus* (Steindachner 1876)	K^b (1)	
	tuyra (Meek and Hildebrand 1923)	M^f (15)	Species is known to occur in Pacific estuaries and lower reaches of rivers (Fischer et al. 1995). The species has been found in Lake Alajuela and Lake Gatun in the Caribbean watershed, in the latter even being reproductively active (D. Sharpe, pers. comm., December, 2015), but no official report of occurrences of *C. tuyra* on the Atlantic side of Panama has been made. Our sample of *C. tuyra* originates from Puente del Río Chagres, located between the Panama Canal and Lake Alajuela, approx. 59 km inland (distances were calculated following meanders with Google Earth). Therefore, the sampled population can be considered as true freshwater inhabitants

Table 1. Continued.

Genus	Species	Localities[1]	Notes
	fuerthii (Steindachner 1876)	L^b (1), J^b (4)	
	sp. indet.	H^b (4)	All 4 sequenced individuals had identical *ATPase 8/6* sequence that is not present in the reference dataset, but had sequence affinity to *Cathorops*
	**nuchalis* (Günther 1864)	A^f (12)	**ATPase 8/6* sequence is identical for both species. Shape data for both species were merged in the phylogenetic analysis as the phylogenetic tree lacks the resolution of both species. They were treated as individual species in PCA (Fig. 3) and CVA (Fig. 5)
	**wayuu* (Betancur-R et al. 2012)	B^b (20), C^b (14), D^m (6), F^m (9)	

[1]Capital letter refers to sampling locality illustrated in Figure 1; small superscript letter refers to habitat characteristic – f: fresh, b: brackish, and m: salt water; and numbers in brackets refer to number of sampled individuals in that specific locality.

Opercle shape changes within Ariidae according to habitat

To assess the shape changes that occur among specimens from the sampled habitats (freshwater, brackish, and marine), a CVA was conducted on the specified groups. The first two CVs explained 100% of the observed variation among freshwater, brackish, and marine samples (Fig. 5A). CV1 separated the three habitats, with the marine group exhibiting the most negative values, the brackish group having values smaller than zero, and the freshwater group having positive values. CV2 separated the brackish group from the marine and freshwater group with the first having values larger than zero and the latter having values smaller than zero. All habitats were distinct in shape as measured by both Procrustes and Mahalanobis distances (Table 2). The most pronounced group differences were found in the pairwise comparison of freshwater and marine shape data. The underlying opercle shape change from marine to freshwater habitat (Fig. 5B, CV1) was reflected in a transformation of a rather triangular opercle shape without any distinguishable features to a more complex shape. A ventral protuberance on the posterior edge makes the overall shape of the ventral surface blunter. However, the meeting point of the posterior and anterior edge becomes more pointed. The ventral part of the anterior edge becomes progressively more sigmoidally curved with a deeper notch forming right behind the articular process. The shape change in CV2 was less pronounced although 32.5% of the observed variation is captured along that axis.

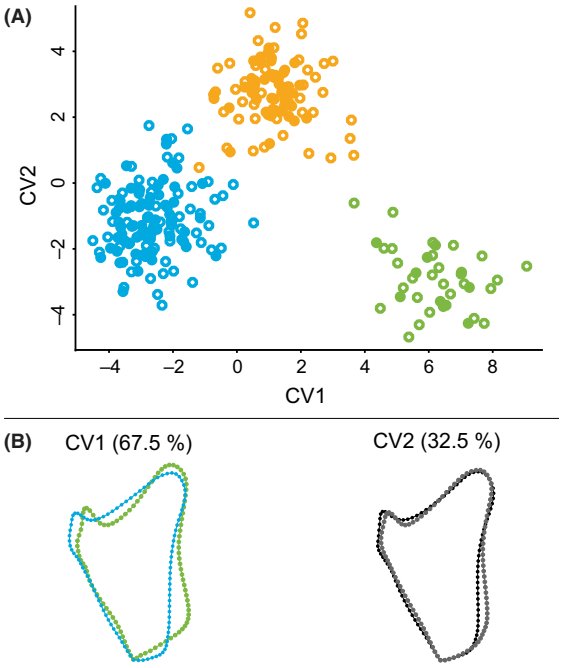


Figure 5. Opercle shape variation in ariid catfish species from marine, brackish, and freshwater habitat. (A) The upper figure shows a scatterplot of individual scores from canonical variate analysis (CVA) comparing ariid catfishes from freshwater (●), brackish (●), and marine (●) habitat for opercle shape. The first two canonical variates (CVs) capture the entire variance (100%) observed between the three groups. CV1 separates them from negative to positive values, from marine, through brackish, to freshwater species, respectively. (B) Opercle shape changes are presented for CV1 and CV2. For CV1, scale factors −4 (blue, representing marine opercle shape) and +8 (green, representing freshwater opercle shape) were applied, and for CV2, −4 (black) and +4 (gray).

Discussion and Conclusion

Species identification using a mtDNA marker

Molecular species identification led to three main conclusions. First, taxonomic identification using *ATPase 8/6* demonstrates that the taxonomy of tropical American ariid catfishes is still unclear. Based on *ATPase 8/6* haplotypes, *B. pinnimaculatus* seems to contain two distinct taxa. The haplotypes correlate with morphological

Table 2. Distance matrices for opercle Procrustes landmark data derived from canonical variate analysis (CVA) of groups defined by habitat occupation.

	f	b	m
f	–	7.5831 (<0.0001)	9.0013 (<0.0001)
b	0.0270 (0.0120)	–	5.4069 (<0.0001)
m	0.0642 (<0.0001)	0.0542 (<0.0001)	–

P-values (shown in parenthesis) calculated by 1000 random permutations per test to determine statistical significance of differences between pairs of habitats. f: freshwater; b: brackish; m: marine. Above diagonal is Mahalanobis distances; below diagonal is Procrustes distances.

features of neurocrania that diverged from the description of *B. pinnimaculatus* (Cooke and Jiménez 2008a), one featuring typical hyperossified frontals and one lacking the hyperossified frontals. This is also suggested by subsequent inspection of opercle shape clusters in morphospace. Additional sampling of eleven more individuals from the eastern part of the Gulf of Panama and subsequent maceration of the skulls confirmed our suspicion that the two morphotypes might rather co-occur than distribute separately along the shore and that the absence of the hyperossification of the frontals is not a rare event. It remains to be validated whether these two morphs represent separate species.

Four individuals from Rio San Pedro, PA, with affinity to *Cathorops* could not be assigned to any known *ATPase 8/6* sequence and remain unidentified. *Notarius biffi*, the distribution of which was believed to be restricted from El Salvador to Costa Rica (Betancur-R and Acero 2004), was confirmed by molecular analyses to be present south of El Salvador as inferred from distribution records (Robertson and Allen 2015). We therefore strongly recommend the inclusion of molecular-based identification in similar studies that focus on the taxonomy of ariid catfishes and the description of new species. Wrong assignment of a specimen to a different species biases possible results and conclusions drawn from them.

Second, the DNA sequencing was helpful in showing that the Paraguaná Peninsula, VE, might be a geographic dispersal barrier to populations of marine *S. herzbergii*, as manifested in distinct *ATPase 8/6* haplotypes. Such a geographic barrier was demonstrated for freshwater species from the Venezuelan coast (Rodriguez-Olarte et al. 2009), but coastal marine species were so far not included in the biogeographic study of species occurrences. Coastal marine species might face similar migration barriers as freshwater species. This is important when considering possible biogeographic scenarios of the initial dispersal of Ariidae in deep time, which is an occurrence throughout the Tethys Sea for this "family" (Betancur-R et al. 2007). So far the molecular within- and between-species variation has not been assessed in Ariidae, and species from one locality are automatically considered being identical to species in another locality.

Third, we found four recent cases of freshwater invasions where *ATPase 8/6* sequences are still identical between marine/brackish and freshwater populations, as is the case in *C. tuyra*, *N. cookei*, possibly *S. parkeri*, and in the sister species *C. wayuu* and *C. nuchalis*. We note that mitochondrial DNA might fail to detect hybridization or introgression events, and further nuclear-based evidence is needed to evaluate whether incipient speciation is going on.

Cathorops. tuyra has migrated from the Pacific side of Panama, where it inhabits Pacific estuaries and lowland rivers, to the Atlantic watershed of Panama since the termination of the Panama Canal in 1914. It has managed to get through the lock gates of the Panama Canal, which ostensibly were an artificial dispersal barrier, reaching Gatún Lake via the Chagres River, where it now thrives and breeds. It has also been recorded recently in artificial Alajuela Lake, which was formed by damming the River Chagres further upstream in 1934 (D. Sharpe, pers. comm., December 2015). This represents a distance of at least 50 km from the Pacific exit of the Panama Canal. The Chagres River flows into the Atlantic at Fuerte de San Lorenzo, 60 km from our collection station on the Rio Chagres Bridge on the Trans-Isthmian Highway. On the Pacific watershed, A. Vega (pers. comm., December, 2015) recorded *C. tuyra* as far inland as the Gatún River, an affluent of the Santa Maria River, at the bridge, near San Francisco, approx. 95 km from the Santa Maria marine exit.

Notarius cookei was captured in freshwater at Tierra Hueca (76 km inland) and La Raya (69 km) along with *C. tuyra* (A. Vega, pers. comm., December, 2015) (distances were calculated following meanders of the Santa Maria River with Google Earth). In the archaeological record, this species has only been recorded at the boundary of marine influence where surface water has been shown to be fresh (in middens radiocarbon-dated between 1900 and 1500 radiocarbon years ago at Sitio Sierra, now 13 km from the sea) (Cooke and Jiménez 2008b). A priority for the future is to locate more dated refuse dumps with fish remains further inland in order to clarify *N. cookei's* dispersal pattern through time.

Cathorops wayuu (brackish, marine) from the Venezuelan coast and *C. nuchalis* (freshwater) from Maracaibo Lake, VE, are described as being two separate species based on morphology. Their *ATPase 8/6* sequences are identical, though. This species pair might represent an example of freshwater adaptation that is older than in the examples described above, as a disparate morphology has been already noted (Marceniuk et al. 2012).

The identification of *S. parkeri* in the Orinoco River, 320 km inland and outside its known habitat range (Betancur-R et al. 2008), was unexpected; whether a freshwater population has been established remains to be proven. However, we cannot exclude the possibility that *S. parkeri* was brought from the coast to the fish market although it seems unlikely. No coastal fishes are sold at local markets in the amazon region of Venezuela.

Those four cases are ideal for applying high-throughput DNA sequencing targeting the marine/brackish and freshwater populations and species to gain insights into the molecular changes associated with adaptations to freshwater environments.

Geometric morphometrics

Patterns of opercle shape variation among ariid species

The present study aimed to identify patterns in ariid opercle shape variation along a major habitat transition from sea to freshwater. Three examples of opercle shape divergences are known so far from the literature: (1) divergence after freshwater invasion in Alaskan *G. aculeatus* populations (Kimmel et al. 2008) resulting in the same shape optima in different lakes, (2) decaying phylogenetic signal in major clades with opercle shape divergence along benthic–pelagic axis in subfamilies in Antarctic notothenioids (Wilson et al. 2013a), and (3) adaptation to feeding ecology in Lake Tanganyika cichlids (Wilson et al. 2015).

In this study, we found the same pattern of shape evolution (anterior–posterior stretching and dorsal–ventral compression) of the opercle along major axes of variance as was identified in previous studies. Additionally, we found strong phylogenetic structure in the opercle shape data of the different ariid species, meaning that the closer species were related, the more similar their opercles were, conforming to the Brownian model of evolution. Ariid genera formed separate cohorts in morphospace (Fig. 3). These results underline the importance of taking into consideration phylogenetic relationships in the study of ariid opercle shape. Nonetheless, similar opercle shapes were found when comparing species from one habitat to species from another habitat. This implies the relevance of habitat on opercle shape to a certain degree. We did not detect any subclade divergence as in Antarctic notothenioids.

The three examples from the literature and our examples from Ariidae differ in several aspects, but mainly by clade age, species richness, and the kind of habitat change. This makes each example rather unique in its evolutionary history. *G. aculeatus* (stickleback) freshwater populations represent an example of repeated evolution resulting in the same shape optima in the different lakes (Arif et al. 2009; Kimmel et al. 2012a), which formed after the last ice age (Bell and Foster 1994), and are sometimes as young as 1983 in the case of Loberg Lake, USA (Arif et al. 2009). Antarctic notothenioids and Ariidae are both radiations that are millions of years old, 17.1–10 ma (Colombo et al. 2015) and 70–40 ma (Betancur-R 2010; Betancur-R et al. 2012), respectively. The radiation of Ariidae was limited by competition (Betancur-R et al. 2012), which might have constrained shape evolution. Antarctic notothenioids on the other hand, encountered only weak competition (Wilson et al. 2013a), which might have facilitated the decay of phylogenetic signal and allowed for divergence along the benthic–pelagic axis in subfamilies. Finally, Lake Tanganyika cichlids are among the most species-rich vertebrate adaptive radiations (Salzburger et al. 2014). They seem to have radiated into different trophic niches first and only diversified later along the benthic–limnetic axis (Muschick et al. 2014). This inference is in agreement with the finding that opercle shape correlates with gut length as a proxy for feeding mode rather than habitat (Wilson et al. 2015). Our study of the opercle bone in Ariidae allowed us to identify a different pattern of shape evolution compared to the other studied clades. Additionally, the importance of the opercle bone in Ariidae lies in its taxonomic value and the possibility for comparison with the fossil record as was shown in extinct marine species of the garlike *Saurichthys* (Wilson et al. 2013b).

Opercle shape changes within Ariidae according to habitat

Although opercle shapes were more similar among closely related species, opercles from freshwater and marine species are significantly distinct in multivariate analyses. For the creation of the phylomorphospace only species were included that were present in the phylogenetic tree, merging *C. tuyra* brackish and freshwater populations, *S. dowii* brackish and freshwater populations, and *C. nuchalis* (freshwater) and *C. wayuu* (brackish and marine) as one species. Therefore, any possible deviating signal from those contrasting habitat populations is lost in the phylomorphospace plot. However, when treated separately such as in the CVA, the three habitat groups were clearly separated, manifesting in a shape change from a rather triangular opercle shape in marine species to a more complex shape in freshwater species. However, we could not demonstrate whether that signal holds when shape data were corrected for phylogenetic dependence of shape values. In order to truly account for the phylogenetic component in the shape data, in a multivariate analysis such

as phylogenetic generalized least squares (pGLS) or phylogenetic ANOVA, a time-calibrated phylogenetic tree would be needed that contains all analyzed species and populations originating from the different habitats.

To conclude, opercle shape is largely conserved across phylogeny in tropical American Ariidae, with closer-related species being characterized by more similar opercle shapes. This fact validates the application of opercle shape for taxonomic identification of fossil bones. On the other hand, opercle shape in ariid catfishes appears to reflect adaptations along the marine–freshwater macrohabitat axis. Sticklebacks, Antarctic notothenioids, Lake Tanganyika cichlids, and tropical American Ariidae exhibit the same dorsal–ventral compression and anterior–posterior stretching along major axes of variance, suggesting this to be a general trend in teleosts.

Acknowledgments

We thank Aureliano Valencia, Máximo Jiménez, and Carlos Jaramillo from the Smithsonian Institute, Panama, and Gilson Rivas from the Universidad del Zulia, Venezuela, Cathy Villalba, Venezuela, and Alexandra Wegmann, University of Zurich, for support in the field with logistics and rapid preliminary identification of collected taxa; Linda Frey, University of Zurich, for help with digitizing the opercles; and Roxana Segundo from the Smithsonian Institute, Panama, for assistance with the calculations of distances in Panamanian rivers. Special thanks are accorded to Jorge Carrillo-Briceño whose support and relentless negotiations made the fieldwork possible. We also thank Ricardo Betancur-R for sharing his taxonomic expertise on Ariidae with us and Melanie Cristescu and three anonymous reviewers for valuable feedback on the manuscript. This project was supported by the Swiss National Fund Sinergia project granted to MRSV and WS (CRSII3-136293). MS was funded by Forschungskredit of the University of Zurich, grant no. FK-15-092. Research conducted at the Smithsonian Tropical Research Institute (STRI), Panama, and in Venezuela was covered by Collecting permit No. 59, 2014 and permit No. 1, 2014, respectively.

Conflict of Interest

None declared.

References

Acero, P. A. 2002. Order Siluriformes - Ariidae. Pp. 831–852 *in* K. Carpenter, ed. The living marine resources of the western central Atlantic, FAO, Rome.

Acero, P. A., and R. Betancur-R. 2002. *Arius cookei*, a new species of ariid catfish from the tropical American Pacific. Aqua, J. Ichthyol. Aquatic Biol., 5:133–138.

Acero, P. A., and R. Betancur-R. 2007. Monophyly, affinities, and subfamilial clades of sea catfishes (Siluriformes: Ariidae). Ichthyol. Explor. Freshw. 18:133–143.

Adams, D. C., and E. Otarola-Castillo. 2013. geomorph: an R package for the collection and analysis of geometric morphometric shape data. Methods Ecol. Evol. 4: 393–399.

Aguilera, O., and D. R. de Aguilera. 2004. Amphi-American Neogene sea catfishes (Siluriformes, Ariidae) from Northern South America. Spec. Pap. Paleontol. 71:29–48.

Aguilera, O., J. G. Lundberg, J. Birindelli, M. Sabaj Pérez, C. Jaramillo, and M. R. Sánchez-Villagra. 2013. Palaeontological evidence for the last temporal occurrence of the ancient Western amazonian river outflow into the Caribbean. PLoS ONE 8:e76202.

Alfaro, M. E., F. Santini, C. Brock, H. Alamillo, A. Dornburg, D. L. Rabosky, et al. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proc. Natl Acad. Sci. USA 106:13410–13414.

Arif, S., W. E. Aguirre, and M. A. Bell. 2009. Evolutionary diversification of opercle shape in Cook Inlet threespine stickleback. Biol. J. Linn. Soc. 97:832–844.

Bell, M. A., and S. A. Foster. 1994. Introduction to the evolutionary biology of the threespine stickleback. Pp. 1–27 *in* M. A. Bell, S. A. Foster, eds. The evolutionary biology of the threespine stickleback. Oxford Univ. Press, Oxford, UK.

Betancur-R, R. 2009. Molecular phylogenetics and evolutionary history of ariid catfishes revisited: a comprehensive sampling. BMC Evol. Biol. 9:175.

Betancur-R, R. 2010. Molecular phylogenetics supports multiple evolutionary transitions from marine to freshwater habitats in ariid catfishes. Mol. Phylogenet. Evol. 55:249–258.

Betancur-R, R., and A. Acero. 2004. Description of *Notarius biffi* n. sp. and redescription of *N. insculptus* (Jordan and Gilbert) (Siluriformes: Ariidae) from the eastern Pacific, with evidence of monophyly and limits of *Notarius*. Zootaxa, 703:1–20.

Betancur-R, R., and P. A. Acero. 2005. Description of *Cathorops mapale*, a new species of sea catfish (Siluriformes: Ariidae) from the Colombian Caribbean, based on morphological and mitochondrial evidence. Zootaxa 1045:45–60.

Betancur-R, R., P. A. Acero, E. Bermingham, and R. Cooke. 2007. Systematics and biogeography of New World sea catfishes (Siluriformes: Ariidae) as inferred from mitochondrial, nuclear, and morphological evidence. Mol. Phylogenet. Evol. 45:339–357.

Betancur-R, R., A. P. Marceniuk, and P. Béarez. 2008. Taxonomic status and redescription of the Gillbacker sea catfish (Siluriformes: Ariidae: *Sciades parkeri*). Copeia 4:827–834.

Betancur-R, R., G. Orti, A. M. Stein, A. P. Marceniuk, and R. A. Pyron. 2012. Apparent signal of competition limiting diversification after ecological transitions from marine to freshwater habitats. Ecol. Lett. 15:822–830.

Bloom, D. D., J. T. Weir, K. R. Piller, and N. R. Lovejoy. 2013. Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analyses and molecular phylogenetics of New World silversides (Atherinopsidae). Evolution 67:2040–2057.

Bookstein, F. L. 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Med. Image Anal. 1:225–243.

Bookstein, F. L., K. Schafer, H. Prossinger, H. Seidler, M. Fieder, G. Stringer, et al. 1999. Comparing frontal cranial profiles in archaic and modern Homo by morphometric analysis. Anat. Rec. 257:217–224.

de Carvalho, M. R., and J. D. McEachran. 2003. Family Carcharhinidae (requiem sharks). Pp. 13–16 *in* R. E. Reis, S. O. Kullander, C. J. Ferraris Jr, eds. Check list of the freshwater fishes of South and Central America. ediPUCRS, Porto Alegre.

Chen, W., M. Al-Husaini, M. Beech, K. Al-enezi, S. Rajab, and H. Husain. 2011. Discriminant analysis as a tool to identify catfish (Ariidae) species of the excavated archaeological otoliths. Environ. Biol. Fish 90:287–299.

Colombo, M., M. Damerau, R. Hanel, W. Salzburger, and M. Matschiner. 2015. Diversity and disparity through time in the adaptive radiation of Antarctic notothenioid fishes. J. Evol. Biol. 28:376–394.

Cooke, R., and M. Jiménez. 2008a. Marine catfish (Ariidae) of the tropical eastern pacific: an update emphasising taxonomy, zoogeography, and interpretation of pre-columbian fishing practices. In Bearez, Philippe, Grouard, Sandrine Clavel, Benoît, Archéologie du Poisson 30 Ans d'Archéo-Ichtyologie au CNRS Hommage aux Trav Jean Desse Nathalie Desse-Berset Actes des Rencontres, 18-20 Octobre 2007/XXVIIIe Rencontres Int d'. Éditions A. Edited by Bearez P, Grouard S, Clavel B. Antibes: 161–179.

Cooke, R., and M. Jiménez. 2008b. Pre-Columbian use of freshwater fish in the Santa Maria Biogeographical Province, Panama. Quat. Int. 185:46–58.

Cuvier, G., and A. Valenciennes. 1840. Histoire Naturelle des Poissons. Tome Quinzième. Suite du Livre Dix-septième. Siluroïdes. Ch. Pitois & Ve Levrault, Paris & Strasbourg.

Davis, A. M., P. J. Unmack, B. J. Pusey, J. B. Johnson, and R. G. Pearson. 2012. Marine-freshwater transitions are associated with the evolution of dietary diversification in terapontid grunters (Teleostei: Terapontidae). J. Evol. Biol. 25:1163–1179.

Diogo, R. 2004. Phylogeny, origin and biogeography of catfishes: support for a Pangean origin of "modern teleosts" and reexamination of some Mesozoic Pangean connections between the Gondwanan and Laurasian supercontinents. Anim. Biol. 54:331–351.

Drake, A. G., and C. P. Klingenberg. 2008. The pace of morphological change: historical transformation of skull shape in St Bernard dogs. Proc. Biol. Sci. 275:71–76.

Fischer, W., F. Krupp, W. Schneider, C. Sommer, K. E. Carpenter, and V. H. Niem. 1995. Guia FAO para la identificacion de especies para los fines de la pesca, pacifico centro-oriental Volumen II. Vertebrados - Parte 1. FAO, Roma.

Gill, T. N. 1863. Descriptive enumeration of a collection of fishes from the western coast of Central America, presented to the Smithsonian Institution by Captain John M. Dow. Proc. Acad. Nat. Sci. Philadelphia 15:162–174.

Günther, A. 1864. Catalogue of the Fishes in the British Museum, vol. 5.– Catalogue of the Physostomi, Containing the Families Siluridae, Characinidae, Haplochitonidae, Sternoptychidae, Scopelidae, Stomiatidae in the Collection of the British Museum. Trustees, London, p. 147.

Gunz, P., P. Mitteroecker, and F. L. Bookstein. 2005. Semilandmarks in three dimensions. Pp. 73–98 *in* D. E. Slice, ed. Modern morphometrics in physical anthropology. Kluwer Academic/Plenum, New York, NY.

Hoorn, B. C., and C. Hoorn. 2006. The birth of the mighty Amazon. Sci. Am. May:52–59.

Katoh, K., and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30:772–780.

Kimmel, C. B., B. Ullmann, C. Walker, C. Wilson, M. Currey, P. C. Phillips, et al. 2005. Evolution and development of facial bone morphology in threespine sticklebacks. Proc. Natl Acad. Sci. USA 102:5791–5796.

Kimmel, C. B., W. Aguirre, B. Ullmann, M. Currez, and W. Cresko. 2008. Allometric change accompanies opercular shape evolution in Alaskan threespine sticklebacks. Behaviour 145:669–691.

Kimmel, C. B., A. DeLaurier, B. Ullmann, J. Dowd, and M. McFadden. 2010. Modes of developmental outgrowth and shaping of a craniofacial bone in zebrafish. PLoS ONE 5: e9475.

Kimmel, C. B., W. A. Cresko, P. C. Phillips, B. Ullmann, M. Currey, F. von Hippel, et al. 2012a. Independent axes of genetic variation and parallel evolutionary divergence of opercle bone shape in threespine stickleback. Evolution 66:419–434.

Kimmel, C. B., P. A. Hohenlohe, B. Ullmann, M. Currey, and W. A. Cresko. 2012b. Developmental dissociation in morphological evolution of the stickleback opercle. Evol. Dev. 14:326–337.

Klingenberg, C. P. 2011. MorphoJ: an integrated software package for geometric morphometrics. Mol. Ecol. Resour. 11:353–357.

Klingenberg, C. P., and N. A. Gidaszewski. 2010. Testing and quantifying phylogenetic signals and homoplasy in morphometric data. Syst. Biol. 59:245–261.

Klingenberg, C. P., M. Barluenga, and A. Meyer. 2002. Shape analysis of symmetric structures: quantifying variation among individuals and asymmetry. Evolution 56:1909–1920.

Lasso, C. A., and P. Sánchez-Duarte. 2011. Los peces del delta del Orinoco. Diversidad, bioecología, uso y conservación. Fundación La Salle de Ciencias Naturales y Chevron, Caracas, Venezuela.

Lovejoy, N. R., J. S. Albert, and W. G. R. Crampton. 2006. Miocene marine incursions and marine/freshwater transitions: evidence from Neotropical fishes. J. South Am. Earth Sci. 21:5–13.

Marceniuk, A. P. 2007. Revalidação de *Cathorops arenatus* e *Cathorops agassizii* (Siluriformes, Ariidae), bagres marinhos das regiões norte e nordeste da América do Sul. Iheringia, Série Zool. 97:360–375.

Marceniuk, A. P., R. Betancur-R, P. A. Acero, and J. Muriel-Cunha. 2012. Review of the Genus *Cathorops* (Siluriformes: Ariidae) from the Caribbean and Atlantic South America, with Description of a New Species. Copeia 2012:77–97.

Martins, E. P., and T. F. Hansen. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. Am. Nat. 149:646–667.

Meek, S. E., and S. F. Hildebrand. 1923. The marine fishes of Panama. Part I. Field Mus. Nat. Hist., Zool. Ser. 15:1–330.

Muschick, M., P. Nosil, M. Roesti, M. T. Dittmann, L. Harmon, and W. Salzburger. 2014. Testing the stages model in the adaptive radiation of cichlid fishes in East African Lake Tanganyika Testing the stages model in the adaptive radiation of cichlid fishes in East African Lake Tanganyika. Proc. R. Soc. B 281:20140605.

Nakatani, M., M. Miya, K. Mabuchi, K. Saitoh, and M. Nishida. 2011. Evolutionary history of Otophysi (Teleostei), a major clade of the modern freshwater fishes: Pangaean origin and Mesozoic radiation. BMC Evol. Biol. 11:177.

Orme, D., R. Freckleton, G. Thomas, T. Petzoldt, S. Fritz, N. Isaac, et al. 2013. caper: Comparative Analyses of Phylogenetics and Evolution in R. R package version 0.5.2.

Robertson, D. R., and G. R. Allen. 2015. Shorefishes of the Tropical Eastern Pacific: online information system. Version 2.0 Smithsonian Tropical Research Institute, Balboa, Panamá. http://biogeodb.stri.si.edu/sftep, assessed May 8, 2016.

Rodriguez-Olarte, D., D. C. Taphorn, and J. Lobon-Cervia. 2009. Patterns of freshwater fishes of the caribbean versant of Venezuela. Int. Rev. Hydrobiol. 94:67–90.

Rohlf, F. J. 2010. tpsRelw: relative warps analysis. Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, NY.

Rohlf, F. 2013. tpsDig. Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, NY.

Salzburger, W., B. Van Bocxlaer, and A. S. Cohen. 2014. Ecology and evolution of the African Great Lakes and Their Faunas. Annu. Rev. Ecol. Evol. Syst. 45:519–545.

Scheyer, T. M., L. Schmid, H. Furrer, and M. R. Sánchez-Villagra. 2014. An assessment of age determination in fossil fish: the case of the opercula in the Mesozoic actinopterygian Saurichthys. Swiss J. Palaeontol. 133:243–257.

Stamatakis, A. 2006. RAxML-VI-HPC : maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

Steindachner, F. 1876. Ichthyologische Beiträge, IV. Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, Mathematisch-Naturwissenschaftlichen Classe, Wien, Abt. 1. Botanik, Zoologie, Anatomie, Geologie und Paläontologie 72:551–616. (pls. 1–13).

Sullivan, J. P., J. G. Lundberg, and M. Hardman. 2006. A phylogenetic analysis of the major groups of catfishes (Teleostei: Siluriformes) using rag1 and rag2 nuclear gene sequences. Mol. Phylogenet. Evol. 41:636–662.

Teugels, G. G. 1996. Taxonomy, phylogeny and biogeography of catfishes (Ostariophysi, Siluroidei): an overview. Aquatic Living Resources, 9, Pp. 9–34.

Traill, T. S. 1832. Description of a *Silurus*, known in Demerara by the name of Gilbacke, more properly Geelbuik. Mem. Wernerian Nat. Hist. Soc. 6:377–380.

Wilson, L. A. B., M. Colombo, R. Hanel, W. Salzburger, and M. R. Sánchez-Villagra. 2013a. Ecomorphological disparity in an adaptive radiation: opercular bone shape and stable isotopes in Antarctic icefishes. Ecol. Evol. 3:3166–3182.

Wilson, L. A. B., H. Furrer, R. Stockar, and M. R. Sánchez-Villagra. 2013b. A quantitative evaluation of evolutionary patterns in opercle bone shape in *Saurichthys* (Actinopterygii: Saurichthyidae). Palaeontology 56:901–915.

Wilson, L. A. B., M. Colombo, M. R. Sánchez-Villagra, and W. Salzburger. 2015. Evolution of opercle shape in cichlid fishes from Lake Tanganyika - adaptive trait interactions in extant and extinct species flocks. Sci. Rep. 5:16909.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Appendix S1. Extended methods.
Appendix S2. Extended results.
Table S1. Sampling Locations and definition of habitat of species used in this study.
Table S2. Number of individuals per species used in PCA (Fig. 3a) and CVA (Fig. 5), and per habitat in CVA.
Table S3. Unique identifiers, sampling locality, and species names for all 263 individuals used in this study.
Figure S1. Scatterplots of the first three principal components (PCs) of individual ariid opercle shapes ($N = 263$) from 21 species belonging to the genera *Bagre* (●), *Sciades* (●), *Cathorops* (●), *Notarius* (●) and *Ariopsis* (●).

Chapter 1

Evolution of opercle bone shape along a macrohabitat gradient: species identification using mtDNA and geometric morphometric analyses in Neotropical sea catfishes (Ariidae)

Supplementary material

Appendix S1. Extended methods.

Details on *ATPase 8/6* amplification and sequencing conditions.

For amplification and sequencing of mitochondrial *ATPase 8/6* primers 8.2 L8331 5'-AAAGCRTYRGCCTTTTAAGC-3' and 3.2 H9236 5'-GTTAGTGGTCAKGGGCTTGGRTC-3' were used (Betancur-R. *et al.*, 2007). *ATPase 8/6* was amplified using REDTaq® DNA polymerase (Sigma-Aldrich). Cycling conditions were as follows: 1 cycle: 94 °C 2 min; 25 cycles: 94 °C 30 s, 54 °C 40 s and 72 °C 1 min; 1 cycle: 72 °C 10 min. PCR-products were enzymatically purified by using the ExoSAP-IT™ PCR Clean-Up Kit (GE Healthcare). The templates were sequenced on an Applied Biosystems 3130xl Genetic Analyzer (Life Technologies) using the Big Dye® Terminator Ready Reaction Mix 3.1 (Applied Biosystems). Sequencing conditions: 1 cycle: 96 °C 1 min; 25 cycles: 96 °C 15 s, 51°C 15 s, 60 °C 4 min.

Appendix S2. Extended results.

Species identification using the mtDNA marker *ATPase 8/6*.

Sciades. *S. proops* (Valenciennes, 1840) from the Gulf of Venezuela (marine) and Puerto La Cruz (marine), and *S. herzbergii* (Bloch, 1784) from the Gulf of Venezuela (marine) and Clarines (brackish), VE, clustered according to species, with both species exhibiting two separate clusters of populations according to geographic origin of samples. *S. dowii* (Gill, 1863) specimens from the mouth of Río Hato, Puerto Caimito (brackish) and Río Santa Maria (freshwater), all from the Pacific side of Panama, were identical in *ATPase 8/6* sequence. Two individuals from the Orinoco River (freshwater) bought at a fish market in Ciudad Bolívar, VE, located 320 km inland, were identified as the marine-brackish *S. parkeri* (Traill, 1832). This species is described as occurring in lower parts of rivers from the Gulf of Paria, VE, to Brazil (Betancur-R. *et al.*, 2008).

Ariopsis. Brackish *A. seemanni* (Günther, 1864) specimens from Río Estero Salado at the Pacific side of Panama and marine *Ariopsis sp. nov.* (Ricardo Betancur-R.,

personal communication, May 2014) specimens from Casaya, Pearl Islands, PA, were confirmed.

Notarius. *N. cookei* (Acero P. & Betancur-R., 2002) is described as being a brackish species (Betancur-R. *et al.*, 2007). Specimens from Santa Maria River 10 km upstream (in freshwater), PA, were confirmed by the *ATPase 8/6* sequence. *N. kessleri* (Steindachner, 1876), sampled in the estuary of Rio San Pedro, near Montijo, and Rio Estero Salado, PA, were confirmed. A single specimen each of *N. biffi* (Betancur-R. & Acero P., 2004) from the Río San Pedro estuary, and *N. planiceps* from Rio Estero Salado, PA, could be confirmed by *ATPase 8/6* sequence. *N. quadriscutis* (Valenciennes, 1840) from Clarines, VE, and *N. grandicassis* (Valenciennes, 1840) specimens from the Gulf of Venezuela formed a separate clade each. A single individual of

Bagre. *B. bagre* (Linnaeus, 1766) from the Gulf of Venezuela (marine) and *B.* aff. *marinus* (marine) from the Gulf of Venezuela and Puerto La Cruz, VE, were confirmed as these species. *B. panamensis* (Gill, 1863) and *B. pinnimaculatus* (Steindachner, 1876) from the Río Estero Salado estuary (brackish), PA (Pacific side) were also confirmed as these species. Samples identified by morphology as *B. pinnimaculatus* from the Gulf of Panama (marine) showed differences in the neurocrania. Several specimens diverged from the morphology of *B. pinnimaculatus* by lacking the typical hyperossification of the frontals. They did not exhibit the phenotype of *B. panamensis*, either, the only other *Bagre* species occurring in the Eastern Pacific. Morphologically they resembled *B. bagre* from the Atlantic. In the ML tree those specimens formed a sister clade to *B. pinnimaculatus*. However, due to their molecular relatedness they were considered being *B. pinnimaculatus* in the multivariate analyses.

Cathorops. *C. hypophthalmus* (Steindachner, 1876) was verified from the Río Estero Salado estuary (brackish). *C. tuyra* (Meek & Hildebrand, 1923) is known to occur in Pacific estuaries and lower reaches of rivers (Fischer *et al.*, 1995). The species has been found in Lake Alajuela and Lake Gatun, in the latter even being reproductively active (Diana Sharpe, personal communication, December, 2015) but no official report of occurrences of *C. tuyra* on the Atlantic side of Panama has been

made. Our sample of *C. tuyra* originates from Puente del Río Chagres, located between the Panama Canal and Lake Alajuela, approx. 59 km inland (distances were calculated following meanders with Google Earth), therefore the sampled population can be considered as true freshwater inhabitants. The *ATPase 8/6* sequences were identical to the reference sequence from specimens collected in brackish water on the Pacific side (in the reference dataset [Betancur-R., 2009]). *C. fuerthii* (Steindachner, 1876) was sampled in the Pacific drainages of Río Parita and Río Hato (brackish), PA, and verified. Four ariid specimens from brackish habitats in the Río San Pedro estuary, PA, were rapidly assigned to *C. steindachneri* (Gilbert & Starks, 1904), *N. cookei*, *C. tuyra* and *C. multiradiatus* (Günther, 1864). However, they all exhibited an identical *ATPase 8/6* sequence that is not present in the reference dataset and are henceforth declared as *C.* sp. indet. *C. nuchalis* (Günther, 1864) was caught in the south of Lake Maracaibo (freshwater) near Puerto Concha, Zulia state, VE. *C. wayuu* (Betancur-R., Acero P. & Marceniuk, 2012) was sampled in the drainage of the lake into the Atlantic at Isla de Toa and Isla de San Carlos (brackish) as well as along the Atlantic coast (marine) as far as Puerto la Cruz, VE. The *ATPase 8/6* sequences of both species clustered together. Shape data for both species was merged in the phylogenetic analysis as the phylogenetic tree lacks the resolution of both species. They were treated as individual species in normal multivariate analyses.

Table S1. Sampling Locations and definition of habitat of species used in this study.

Location	Country	GPS coordinates	Habitat	Species
Lago de Maracaibo / Isla de Toas (B)	Zulia, Venezuela	10°57'9.50"N 71°38'49.54"W	brackish	*C. wayuu**
Lago de Maracaibo /Isla de San Carlos (B)	Zulia, Venezuela	10°59'55.1"N 71°36'19.8"W	brackish	*B. bagre*
Lago de Maracaibo / Puerto Concha (A)	Zulia, Venezuela	9°05'46.0"N 71°42´52"W	freshwater	*C. nuchalis**
Lago de Maracaibo / Guarico (C)	Zulia, Venezuela	10°43'52.0"N 71°31'40.2"W	brackish	*C. wayuu**
Gulf of Venezuela (D)	Falcon, Venezuela	11°14'15.3"N 70°30'53.1" W	marine	*S. proops, S. herzbergii, C. wayuu, B. aff. marinus, B. bagre, N. grandicassis*
Clarines (E)	Anzoategui, Venezuela	10° 3'46.76"N 65°11'5.23"W	brackish	*N. quadriscutis, S. herzbergii, S. proops*
Puerto La Cruz (F)	Anzoategui, Venezuela	10°12'58.63"N 64°38'39.16"W	marine	*C. wayuu*, B. aff. marinus S. proops*
Ciudad Bolivar (G)	Bolívar, Venezuela	8°8'51.46" N 63°32'10.68"W	freshwater	*S. parkeri*
Pearl Islands / Casaya island (P)	Panama	8°34'38.64"N 79°3'3.636" W	marine	*A. nov. sp.*
Puente del Rio Chagres (M)	Panama	9°11'34.66"N 79°39'9.42"W	freshwater	*C. tuyra*
Rio Hato (L)	Panama	8°20'32.4"N 80°09'56.4"W	brackish	*C. fuerthi, S. dowii*
Rio Santa Maria (I)	Panama	8° 6'20.30"N 80°33'16.06"W	freshwater	*N. cookei, S. dowii*
Rio Parita (J)	Panama	8°01'13.69"N 80°27'11.15"W	brackish	*C. fuerthi*
Rio Estero Salado (K)	Panama	8°10'30.324"N 80°29'35.052" W	brackish	*B. pinnimaculatus, B. panamensis, N. planiceps, N. kessleri, C. hypophthalmus, A. seemanni*
Rio San Pedro (H)	Panama	7°50' 59.208"N 81°07' 3.972" W	brackish	*N. kessleri, N. biffi, C.* sp. indet.
Puerto Caimito (N)	Panama	8°52'18.88"N 79°42'32.99"W	marine	*S. dowii*
Gulf of Panama (O)	Panama	8°48'56.55"N 79°22'50.85"W	marine	*B. pinnimaculatus*

Please note that *C. wayuu* und *C. nuchalis* are differentiated here on morphological basis as two separate species although the genetic evidence is missing. Both species together are summarized as *C. sp* in multivariate analyses that depend on a phylogenetic tree. Letters in brackets refer to locations in the sampling map Fig. 1.

Table S2. Number of individuals per species used in PCA (Fig. 3a) and CVA (Fig. 5), and per habitat in CVA.

Species	Freshwater	Brackish	Marine	total
Ans	0	0	9	9
Ase	0	5	0	5
Bba	0	0	5	5
Bma	0	0	19	19
Bpa	0	3	0	3
Bpi	0	1	24	25
Cfu	0	5	0	5
Chy	0	1	0	1
Csp	12	34	15	61
Ctu	15	0	0	15
Cun	0	4	0	4
Nbi	0	1	0	1
Nco	7	0	0	7
Ngr	0	0	11	11
Nke	0	10	0	10
Npl	0	1	0	1
Nqu	0	10	0	10
Sdo	2	3	0	5
She	0	15	31	46
Spa	2	0	0	2
Spr	0	1	17	18
total	**38**	**94**	**131**	**263**

Table S3. Unique identifiers, sampling locality, and species names for all 263 individuals used in this study.

Identifier	Locality	Species abb.	NCBI accession no.
08E2	Pearl Islands / Casaya island, PA	*A.* sp. nov.	KX500399
08E3	Pearl Islands / Casaya island, PA	*A.* sp. nov.	KX500400
08E4	Pearl Islands / Casaya island, PA	*A.* sp. nov.	KX500401
08E5	Pearl Islands / Casaya island, PA	*A.* sp. nov.	KX500402
08E6	Pearl Islands / Casaya island, PA	*A.* sp. nov.	KX500403
08E8	Pearl Islands / Casaya island, PA	*A.* sp. nov.	KX500404
08E9	Pearl Islands / Casaya island, PA	*A.* sp. nov.	KX500405
08F1	Pearl Islands / Casaya island, PA	*A.* sp. nov.	KX500406
08F2	Pearl Islands / Casaya island, PA	*A.* sp. nov.	KX500407
08C6	Rio Estero Salado, PA	*A. seemanni*	KX500408
08C7	Rio Estero Salado, PA	*A. seemanni*	KX500409
08C8	Rio Estero Salado, PA	*A. seemanni*	KX500410
08C9	Rio Estero Salado, PA	*A. seemanni*	KX500411
08D1	Rio Estero Salado, PA	*A. seemanni*	KX500412
04A8	Gulf of Venezuela	*B. marinus*	KX500418
04E9	Gulf of Venezuela	*B. marinus*	KX500419
04F1	Gulf of Venezuela	*B. marinus*	KX500420

Identifier	Locality	Species abb.	NCBI accession no.
04F2	Gulf of Venezuela	*B. marinus*	KX500421
04G5	Gulf of Venezuela	*B. marinus*	KX500422
04G6	Gulf of Venezuela	*B. marinus*	KX500423
04G7	Gulf of Venezuela	*B. marinus*	KX500424
04G8	Gulf of Venezuela	*B. marinus*	KX500425
04G9	Gulf of Venezuela	*B. marinus*	KX500426
04H6	Gulf of Venezuela	*B. marinus*	KX500427
04H7	Gulf of Venezuela	*B. marinus*	KX500428
04H8	Gulf of Venezuela	*B. marinus*	KX500429
05I5	Gulf of Venezuela	*B. marinus*	KX500430
05I6	Gulf of Venezuela	*B. marinus*	KX500431
05I7	Gulf of Venezuela	*B. marinus*	KX500432
06A2	Gulf of Venezuela	*B. marinus*	KX500433
06A6	Puerto La Cruz, VE	*B. marinus*	KX500434
06A7	Puerto La Cruz, VE	*B. marinus*	KX500435
06A8	Puerto La Cruz, VE	*B. marinus*	KX500436
04F5	Gulf of Venezuela	*B. bagre*	KX500413
04F6	Gulf of Venezuela	*B. bagre*	KX500414
04F7	Gulf of Venezuela	*B. bagre*	KX500415
05A8	Lago de Maracaibo /Isla de San Carlos, VE	*B. bagre*	KX500416
05A9	Lago de Maracaibo /Isla de San Carlos, VE	*B. bagre*	KX500417
08B8	Rio Estero Salado, PA	*B. panamensis*	KX500438
08B9	Rio Estero Salado, PA	*B. panamensis*	KX500439
08C1	Rio Estero Salado, PA	*B. panamensis*	KX500440
01A1	Gulf of Panama	*B. pinnimaculatus**	KX500643
01A3	Gulf of Panama	*B. pinnimaculatus**	KX500656
01A5	Gulf of Panama	*B. pinnimaculatus**	KX500644
01A7	Gulf of Panama	*B. pinnimaculatus**	KX500645
01A9	Gulf of Panama	*B. pinnimaculatus**	KX500646
01B2	Gulf of Panama	*B. pinnimaculatus**	KX500647
01B4	Gulf of Panama	*B. pinnimaculatus**	KX500648
01B6	Gulf of Panama	*B. pinnimaculatus*	KX500441
01B8	Gulf of Panama	*B. pinnimaculatus**	KX500657
01C1	Gulf of Panama	*B. pinnimaculatus**	KX500649
01C3	Gulf of Panama	*B. pinnimaculatus*	KX500442
01C5	Gulf of Panama	*B. pinnimaculatus**	KX500655
01C8	Gulf of Panama	*B. pinnimaculatus*	KX500443
01C9	Gulf of Panama	*B. pinnimaculatus*	KX500444
01D2	Gulf of Panama	*B. pinnimaculatus**	KX500658
01D4	Gulf of Panama	*B. pinnimaculatus**	KX500659
01D7	Gulf of Panama	*B. pinnimaculatus**	KX500650
01D8	Gulf of Panama	*B. pinnimaculatus**	KX500651
01E1	Gulf of Panama	*B. pinnimaculatus**	KX500660
01E3	Gulf of Panama	*B. pinnimaculatus**	KX500661
01E5	Gulf of Panama	*B. pinnimaculatus**	KX500652
01E8	Gulf of Panama	*B. pinnimaculatus**	KX500437
01E9	Gulf of Panama	*B. pinnimaculatus**	KX500653
01F2	Gulf of Panama	*B. pinnimaculatus**	KX500654
08B7	Rio Estero Salado, PA	*B. pinnimaculatus*	KX500445

Identifier	Locality	Species abb.	NCBI accession no.
08D3	Rio Parita, PA	*C. fuerthi*	KX500446
08D4	Rio Parita, PA	*C. fuerthi*	KX500447
08D5	Rio Parita, PA	*C. fuerthi*	KX500448
08D6	Rio Parita, PA	*C. fuerthi*	KX500449
08D7	Rio Parita, PA	*C. fuerthi*	KX500450
08C4	Rio Estero Salado, PA	*C. hypophthalmus*	KX500451
01F4	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500453
01F5	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500454
01F6	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500455
01F7	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500456
01G1	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500457
01G2	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500458
01G8	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500459
01H6	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500460
01H9	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500461
07A1	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500462
07A6	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500463
07B2	Lago de Maracaibo / Puerto Concha, VE	*C. nuchalis*	KX500464
08A4	Rio San Pedro, PA	*C.* sp. indet.	KX500452
08A5	Rio San Pedro, PA	*C.* sp. indet.	KX500466
08A6	Rio San Pedro, PA	*C.* sp. indet.	KX500465
08B5	Rio San Pedro, PA	*C.* sp. indet.	KX500542
08F3	Puente del Rio Chagres, PA	*C. tuyra*	KX500467
08F4	Puente del Rio Chagres, PA	*C. tuyra*	KX500468
08F5	Puente del Rio Chagres, PA	*C. tuyra*	KX500469
08F6	Puente del Rio Chagres, PA	*C. tuyra*	KX500470
08F7	Puente del Rio Chagres, PA	*C. tuyra*	KX500471
08F8	Puente del Rio Chagres, PA	*C. tuyra*	KX500472
08F9	Puente del Rio Chagres, PA	*C. tuyra*	KX500473
08G1	Puente del Rio Chagres, PA	*C. tuyra*	KX500474
08G2	Puente del Rio Chagres, PA	*C. tuyra*	KX500475
08G3	Puente del Rio Chagres, PA	*C. tuyra*	KX500476
08G4	Puente del Rio Chagres, PA	*C. tuyra*	KX500477
08G5	Puente del Rio Chagres, PA	*C. tuyra*	KX500478
08G6	Puente del Rio Chagres, PA	*C. tuyra*	KX500479
08G7	Puente del Rio Chagres, PA	*C. tuyra*	KX500480
08G8	Puente del Rio Chagres, PA	*C. tuyra*	KX500481
03G1	Puerto La Cruz, VE	*C. wayuu*	KX500482
03G3	Puerto La Cruz, VE	*C. wayuu*	KX500483
03G5	Puerto La Cruz, VE	*C. wayuu*	KX500484

Identifier	Locality	Species abb.	NCBI accession no.
03G6	Puerto La Cruz, VE	*C. wayuu*	KX500485
03G7	Puerto La Cruz, VE	*C. wayuu*	KX500486
03G8	Puerto La Cruz, VE	*C. wayuu*	KX500487
03G9	Puerto La Cruz, VE	*C. wayuu*	KX500488
03H1	Puerto La Cruz, VE	*C. wayuu*	KX500489
03H2	Puerto La Cruz, VE	*C. wayuu*	KX500490
04A5	Gulf of Venezuela	*C. wayuu*	KX500491
04A6	Gulf of Venezuela	*C. wayuu*	KX500492
04A9	Gulf of Venezuela	*C. wayuu*	KX500493
04B1	Gulf of Venezuela	*C. wayuu*	KX500494
04F3	Gulf of Venezuela	*C. wayuu*	KX500495
04F4	Gulf of Venezuela	*C. wayuu*	KX500496
05B1	Lago de Maracaibo / Guarico	*C. wayuu*	KX500497
05B3	Lago de Maracaibo / Guarico	*C. wayuu*	KX500498
05B4	Lago de Maracaibo / Guarico	*C. wayuu*	KX500499
05B6	Lago de Maracaibo / Guarico	*C. wayuu*	KX500500
05B7	Lago de Maracaibo / Guarico	*C. wayuu*	KX500501
05B8	Lago de Maracaibo / Guarico	*C. wayuu*	KX500502
05B9	Lago de Maracaibo / Guarico	*C. wayuu*	KX500503
05C1	Lago de Maracaibo / Guarico	*C. wayuu*	KX500504
05C2	Lago de Maracaibo / Guarico	*C. wayuu*	KX500505
05C6	Lago de Maracaibo / Guarico	*C. wayuu*	KX500506
05C7	Lago de Maracaibo / Guarico	*C. wayuu*	KX500507
05C8	Lago de Maracaibo / Guarico	*C. wayuu*	KX500508
05D2	Lago de Maracaibo / Guarico	*C. wayuu*	KX500509
05D3	Lago de Maracaibo / Guarico	*C. wayuu*	KX500510
05F3	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500511
05F4	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500512
05F5	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500513
05F6	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500514
05F7	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500515
05F8	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500516
05F9	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500517
05G1	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500518
05G2	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500519
05G4	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500520
05G5	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500521
05G6	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500522
05G7	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500523
05G8	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500524
05G9	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500525
05H1	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500526
05H2	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500527
05H3	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500528
05H4	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500529
05H5	Lago de Maracaibo / Isla de Toas	*C. wayuu*	KX500530
08A9	Rio San Pedro, PA	*N. biffi*	KX500570
08G9	Rio Santa Maria, PA	*N. cookei*	KX500535
08H3	Rio Santa Maria, PA	*N. cookei*	KX500536
08H4	Rio Santa Maria, PA	*N. cookei*	KX500537

Identifier	Locality	Species abb.	NCBI accession no.
08H5	Rio Santa Maria, PA	*N. cookei*	KX500538
08H6	Rio Santa Maria, PA	*N. cookei*	KX500539
08H7	Rio Santa Maria, PA	*N. cookei*	KX500540
08H8	Rio Santa Maria, PA	*N. cookei*	KX500535
04F8	Gulf of Venezuela	*N. grandicassis*	KX500543
04F9	Gulf of Venezuela	*N. grandicassis*	KX500544
04G1	Gulf of Venezuela	*N. grandicassis*	KX500545
04G4	Gulf of Venezuela	*N. grandicassis*	KX500546
04H1	Gulf of Venezuela	*N. grandicassis*	KX500547
04H4	Gulf of Venezuela	*N. grandicassis*	KX500548
04H5	Gulf of Venezuela	*N. grandicassis*	KX500549
05H9	Gulf of Venezuela	*N. grandicassis*	KX500550
05I1	Gulf of Venezuela	*N. grandicassis*	KX500551
05I2	Gulf of Venezuela	*N. grandicassis*	KX500552
08A1	Rio San Pedro, PA	*N. kessleri*	KX500531
08A2	Rio San Pedro, PA	*N. kessleri*	KX500532
08A3	Rio San Pedro, PA	*N. kessleri*	KX500553
08A7	Rio San Pedro, PA	*N. kessleri*	KX500554
08A8	Rio San Pedro, PA	*N. kessleri*	KX500533
08B1	Rio San Pedro, PA	*N. kessleri*	KX500534
08B3	Rio San Pedro, PA	*N. kessleri*	KX500555
08B4	Rio San Pedro, PA	*N. kessleri*	KX500556
08C3	Rio Estero Salado, PA	*N. kessleri*	KX500557
08C5	Rio Estero Salado, PA	*N. kessleri*	KX500559
08C2	Rio Estero Salado, PA	*N. planiceps*	KX500558
03E7	Clarines, VE	*N. quadriscutis*	KX500560
03E8	Clarines, VE	*N. quadriscutis*	KX500561
03E9	Clarines, VE	*N. quadriscutis*	KX500562
03F1	Clarines, VE	*N. quadriscutis*	KX500563
03F3	Clarines, VE	*N. quadriscutis*	KX500564
03F4	Clarines, VE	*N. quadriscutis*	KX500565
03F5	Clarines, VE	*N. quadriscutis*	KX500566
03F7	Clarines, VE	*N. quadriscutis*	KX500567
03F8	Clarines, VE	*N. quadriscutis*	KX500568
03F9	Clarines, VE	*N. quadriscutis*	KX500569
08D8	Puerto Caimito, PA	*S. dowii*	KX500571
08D9	Puerto Caimito, PA	*S. dowii*	KX500572
08E1	Rio Hato, PA	*S. dowii*	KX500573
08H1	Rio Santa Maria, PA	*S. dowii*	KX500574
08H2	Rio Santa Maria, PA	*S. dowii*	KX500575
03A8	Clarines, VE	*S. herzbergii*	KX500576
03C1	Clarines, VE	*S. herzbergii*	KX500577
03C4	Clarines, VE	*S. herzbergii*	KX500578
03C5	Clarines, VE	*S. herzbergii*	KX500579
03C6	Clarines, VE	*S. herzbergii*	KX500580
03C7	Clarines, VE	*S. herzbergii*	KX500581
03C8	Clarines, VE	*S. herzbergii*	KX500582
03C9	Clarines, VE	*S. herzbergii*	KX500583
03D1	Clarines, VE	*S. herzbergii*	KX500584
03D2	Clarines, VE	*S. herzbergii*	KX500585

Identifier	Locality	Species abb.	NCBI accession no.
03D3	Clarines, VE	*S. herzbergii*	KX500586
03D4	Clarines, VE	*S. herzbergii*	KX500587
03D5	Clarines, VE	*S. herzbergii*	KX500588
03D6	Clarines, VE	*S. herzbergii*	KX500589
03D7	Clarines, VE	*S. herzbergii*	KX500590
04A7	Gulf of Venezuela	*S. herzbergii*	KX500591
04B3	Gulf of Venezuela	*S. herzbergii*	KX500592
04B4	Gulf of Venezuela	*S. herzbergii*	KX500593
04B5	Gulf of Venezuela	*S. herzbergii*	KX500594
04B6	Gulf of Venezuela	*S. herzbergii*	KX500595
04B7	Gulf of Venezuela	*S. herzbergii*	KX500596
04B8	Gulf of Venezuela	*S. herzbergii*	KX500597
04B9	Gulf of Venezuela	*S. herzbergii*	KX500598
04C3	Gulf of Venezuela	*S. herzbergii*	KX500599
04C4	Gulf of Venezuela	*S. herzbergii*	KX500600
04C5	Gulf of Venezuela	*S. herzbergii*	KX500601
04C7	Gulf of Venezuela	*S. herzbergii*	KX500602
04C8	Gulf of Venezuela	*S. herzbergii*	KX500603
04C9	Gulf of Venezuela	*S. herzbergii*	KX500604
04D1	Gulf of Venezuela	*S. herzbergii*	KX500605
04D3	Gulf of Venezuela	*S. herzbergii*	KX500606
04D4	Gulf of Venezuela	*S. herzbergii*	KX500607
04D5	Gulf of Venezuela	*S. herzbergii*	KX500608
04D6	Gulf of Venezuela	*S. herzbergii*	KX500609
04D7	Gulf of Venezuela	*S. herzbergii*	KX500610
04D8	Gulf of Venezuela	*S. herzbergii*	KX500611
04D9	Gulf of Venezuela	*S. herzbergii*	KX500612
04E1	Gulf of Venezuela	*S. herzbergii*	KX500613
04E2	Gulf of Venezuela	*S. herzbergii*	KX500614
04E3	Gulf of Venezuela	*S. herzbergii*	KX500615
04E4	Gulf of Venezuela	*S. herzbergii*	KX500616
04E5	Gulf of Venezuela	*S. herzbergii*	KX500617
04E6	Gulf of Venezuela	*S. herzbergii*	KX500618
04E7	Gulf of Venezuela	*S. herzbergii*	KX500619
04E8	Gulf of Venezuela	*S. herzbergii*	KX500620
04H3	Gulf of Venezuela	*S. herzbergii*	KX500621
06B9	Ciudad Bolivar, VE	*S. parkeri*	KX500622
06C1	Ciudad Bolivar, VE	*S. parkeri*	KX500623
03A2	Clarines, VE	*S. proops*	KX500624
04A1	Gulf of Venezuela	*S. proops*	KX500625
04A2	Gulf of Venezuela	*S. proops*	KX500626
04A3	Gulf of Venezuela	*S. proops*	KX500627
04A4	Gulf of Venezuela	*S. proops*	KX500628
04G2	Gulf of Venezuela	*S. proops*	KX500629
04G3	Gulf of Venezuela	*S. proops*	KX500630
04H2	Gulf of Venezuela	*S. proops*	KX500631
04H9	Gulf of Venezuela	*S. proops*	KX500632
05H6	Gulf of Venezuela	*S. proops*	KX500633
05H7	Gulf of Venezuela	*S. proops*	KX500634
05H8	Gulf of Venezuela	*S. proops*	KX500635

Identifier	Locality	Species abb.	NCBI accession no.
05I3	Gulf of Venezuela	*S. proops*	KX500636
05I8	Gulf of Venezuela	*S. proops*	KX500637
05I9	Gulf of Venezuela	*S. proops*	KX500638
06A1	Gulf of Venezuela	*S. proops*	KX500639
06A3	Gulf of Venezuela	*S. proops*	KX500640
06A4	Gulf of Venezuela	*S. proops*	KX500641
06B1	Puerto La Cruz, VE	*S. proops*	KX500642

* *ATPase 8/6* sequence, and neurocranium diverges from *B. pinnimaculatus*

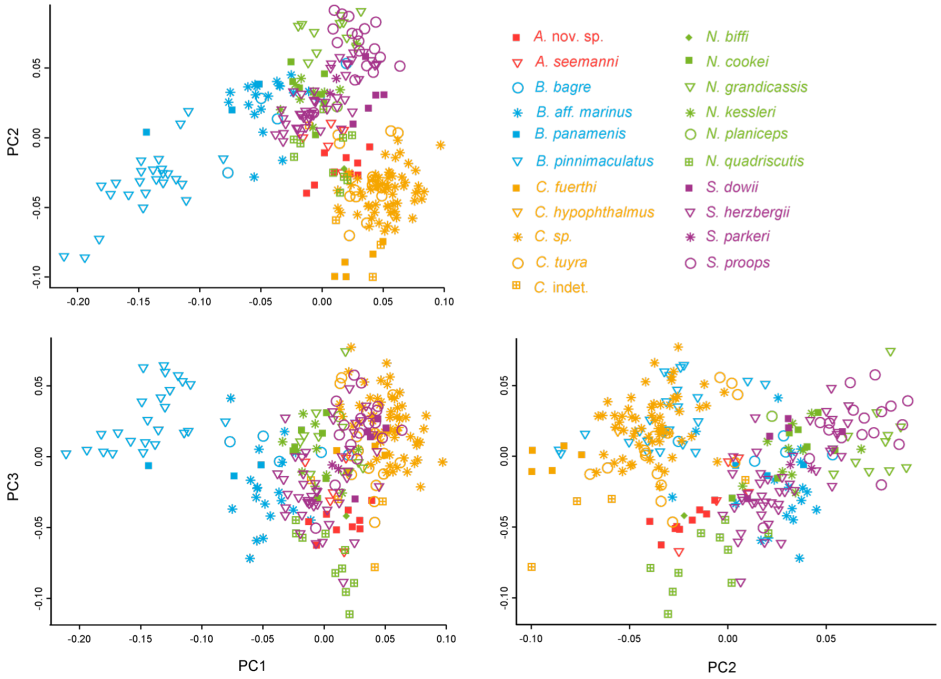


Figure S1. Scatterplots of the first three principal components (PCs) of individual ariid opercle shapes (N=263) from 21 species belonging to the genera *Bagre* (•), *Sciades* (•), *Cathorops* (•), *Notarius* (•) and *Ariopsis* (•). Please note that *C. wayuu* and *C. nuchalis* are pooled here as *C. sp*.

References

Betancur-R., R., & Acero, A. 2004. Description of *Notarius biffi* n. sp. and redescription of *N. insculptus* (Jordan and Gilbert) (Siluriformes: Ariidae) from the eastern Pacific, with evidence of monophyly and limits of *Notarius*. *Zootaxa*, **703**: 1–20.

Betancur-R., R., Acero P., A., Bermingham, E. & Cooke, R. 2007. Systematics and biogeography of New World sea catfishes (Siluriformes: Ariidae) as inferred from mitochondrial, nuclear, and morphological evidence. *Mol. Phylogenet. Evol.* **45**: 339–357.

Betancur-R., R., Acero P., A., Bermingham, E. & Cooke, R. 2007. Systematics and biogeography of New World sea catfishes (Siluriformes: Ariidae) as inferred from mitochondrial, nuclear, and morphological evidence. *Mol. Phylogenet. Evol.* **45**: 339–357.

Betancur-R., R., Marceniuk, A.P. & Béarez, P. 2008. Taxonomic Status and Redescription of the Gillbacker Sea Catfish (Siluriformes: Ariidae: *Sciades parkeri*). *Copeia* **4**: 827–834.

Fischer, W., Krupp, F., Schneider, W., Sommer, C., Carpenter, K.E. & Niem, V.H. 1995. *Guia FAO para la identificacion de especies para los fines de la pesca, pacifico centro-oriental Volumen II. Vertebrados - Parte 1*. FAO, Roma.

Chapter 2

Bayesian divergence-time estimation with genome-wide SNP data of sea catfishes (Ariidae) supports Miocene closure of the Panamanian Isthmus

Bayesian Divergence-Time Estimation with Genome-Wide SNP Data of Sea Catfishes (Ariidae) Supports Miocene Closure of the Panamanian Isthmus

Madlen Stange[1,2], Marcelo R. Sánchez-Villagra[1], Walter Salzburger[2,3], Michael Matschiner[2,3]

[1]*Palaeontological Institute and Museum, University of Zurich, Karl-Schmid-Strasse 4, 8006 Zurich, Switzerland*

[2]*Zoological Institute, University of Basel, 4051 Basel, Switzerland*

[3]*Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biosciences, University of Oslo, 0316 Oslo, Norway*

Corresponding author: Madlen Stange, Palaeontological Institute and Museum, University of Zurich, Karl-Schmid-Strasse 4, 8006 Zurich, Switzerland; E-mail: madlen.stange@pim.uzh.ch; Michael Matschiner, Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland; Email: michaelmatschiner@mac.com

Abstract.—

The closure of the Isthmus of Panama has long been considered to be one of the best defined biogeographic calibration points for molecular divergence-time estimation. However, geological and biological evidence has recently cast doubt on the presumed timing of the initial isthmus closure around 3 Ma but has instead supported the existence of temporary land bridges as early as the Middle or Late Miocene. The biological evidence supporting these earlier land bridges was based either on only few molecular markers or on concatenation of genome-wide sequence data, an approach that is known to result in potentially misleading branch lengths and divergence times, which could compromise the reliability of this evidence. To allow divergence-time estimation with genomic data using the more appropriate multi-species coalescent model, we here develop a new method combining the SNP-based Bayesian species-tree inference of the software SNAPP with a molecular clock model that can be calibrated with fossil or biogeographic constraints. We validate our approach with simulations and use our method to reanalyze genomic data of Neotropical army ants (Dorylinae) that previously supported divergence times of Central and South American populations before the isthmus closure around 3 Ma. Our reanalysis with the multi-species coalescent model shifts all of these divergence times to ages younger than 3 Ma, suggesting that the older estimates supporting the earlier existence of temporary land bridges were artifacts resulting at least partially from the use of concatenation. We further apply our method to a new RAD-sequencing data set of Neotropical sea catfishes (Ariidae) and calibrate their species tree with extensive information from the fossil record. We identify a series of divergences between groups of Caribbean and Pacific sea catfishes around 10 Ma, indicating that processes related to the emergence of the isthmus led to vicariant speciation already in the Late Miocene, millions of years before the final isthmus closure.

(Keywords: Panamanian Isthmus; Central American Seaway; Bayesian inference; phylogeny; molecular clock; fossil record; SNPs; RAD sequencing; teleosts)

The emergence of the Isthmus of Panama had a profound impact on biodiversity in the Western Hemisphere. On land, the isthmus enabled terrestrial animals to migrate between the American continents, which led to massive range expansions and local extinctions during the so-called Great American Biotic Interchange (Woodburne 2010). In the sea, however, the rise of the isthmus created an impermeable barrier between the Caribbean and the Tropical Eastern Pacific (TEP), resulting in the geographic separation of formerly genetically connected marine populations (Lessios 2008). Due to its presumed simultaneous impact on speciation events in numerous terrestrial and marine lineages, the closure of the Isthmus of Panama has been considered one of the best biogeographical calibration points for molecular divergence-time estimation and has been used in several hundreds of phylogenetic studies (Bermingham et al. 1997; Lessios 2008; Bacon et al. 2015a). The precise age of isthmus closure assumed in these studies varies but generally lies between 3.5 Ma (e.g. Donaldson and Wilson Jr 1999) and 2.8 Ma (e.g. Betancur-R. et al. 2012), according to evidence from marine records of isotopes, salinity, and temperature, that all support an age in this range (Jackson and O'Dea 2013; Coates and Stallard 2013; O'Dea et al. 2016).

However, recent research has indicated that the history of the isthmus may be more complex than previously thought and that the isthmus may have closed temporarily millions of years before its final emergence around 3 Ma. The collision between the Panama Arc and the South American plate, which initiated the development of the isthmus, began as early as 25-23 Ma according to geochemical evidence (Farris et al. 2011). As a consequence, the Central American Seaway (CAS), the deep oceanic seaway connecting the West Atlantic and the East Pacific through the Atrato strait, is hypothesized to have narrowed down to a width of 200 km, allowing for continued exchange between the oceans at this time (Farris et al. 2011; Montes et al. 2012). It has been argued that Eocene zircons in Colombian sediments support the existence of Miocene

land bridges and fluvial connections between Panama and South America and thus a closure of the CAS around 15-13 Ma (Montes et al. 2015), however, alternative explanations for the occurrence of these zircons may be possible (O'Dea et al. 2016). Gradual shoaling of the CAS around 11-10 Ma has also been supported by biostratigraphic and paleobathymetric analyses (Coates et al. 2004) as well as seawater isotope records (Sepulchre et al. 2014). On the other hand, a separate analysis of the seawater isotope records indicated that deep-water connections existed until around 7 Ma, followed by mostly uninterrupted shallow-water exhange (Osborne et al. 2014).

While the Atrato strait represented the main connection between the Caribbean and the Pacific throughout most of the Miocene, other passageways existed in the Panama Canal basin (the Panama isthmian strait) and across Nicaragua (the San Carlos strait) (Savin and Douglas 1985). Both of these passageways were likely closed around 8 Ma (and possibly earlier) but reopened around 6 Ma with a depth greater than 200 m, according to evidence from fossil foraminifera (Collins et al. 1996). The last connection between the Caribbean and the Pacific likely closed around 2.8 Ma (O'Dea et al. 2016), but short-lived breachings induced by sea-level fluctuations as late as 2.45 Ma can not be excluded (Groeneveld et al. 2014).

In agreement with the putative existence of earlier land bridges, Miocene dispersal of terrestrial animals between North and South America is well documented in the fossil record. Fossils of a New World monkey, discovered in the Panama Canal basin, demonstrate that primates had arrived on the North American landmass before 20.9 Ma (Bloch et al. 2016). Furthermore, fossils of xenarthran mammals derived from South America (ground sloths, glyptodonts, and pampatheriids) were found in Late Miocene (9-8 Ma) deposits in Florida (Hirschfeld 1968; Laurito and Valerio 2012) and in Early Pliocene (4.8-4.7 Ma) deposits in Mexico (Carranza-Castañeda and Miller 2004; Flynn et al. 2005), and Argentinian fossils of the procyonid carnivore *Cyonasua* provide evidence that

terrestrial mammals had also crossed from North to South America before 7 Ma (Marshall 1988; Bacon et al. 2016). The Argentinian fossils could still be predated by fossils of other mammalian North American immigrants in Late Miocene Amazonian deposits (Campbell et al. 2010; Frailey and Campbell 2012; Prothero et al. 2014), however, their age estimate of 9 Ma may require further confirmation (Carrillo et al. 2015). Dispersal of terrestrial animals is also supported by molecular data. Based on a metaanalysis of phylogenetic data sets, Bacon et al. (2015a,b) reported major increases in migration rates around 10-7 Ma and at 6-5 Ma. In combination with molecular evidence for increased vicariance of marine organisms around 10-9 Ma, the authors concluded that the Isthmus of Panama evolved millions of years earlier than commonly assumed.

Unfortunately, the observed evidence for dispersal of terrestrial organisms is in most cases insufficient for conclusions about the existence of earlier land bridges. This is due to the fact that most of these organisms are members of groups with a known capacity of oceanic dispersal (de Queiroz 2014), in many cases even over far greater distances than the gap remaining between North and South America in the Miocene ($<$ 600 km; Farris et al. 2011). Before their dispersal to the North American landmass in the Early Miocene, primates had already crossed the Atlantic in the Eocene, when they arrived in South America (Kay 2015; Bloch et al. 2016). Many other mammal lineages have proven capable of oversea dispersal, which may be best illustrated by the rich mammalian fauna of Madagascar that is largely derived from Africa even though the two landmasses separated around 120 Ma (Ali and Huber 2010).

As a notable exception without the capacity of oversea dispersal, Winston et al. (2016) recently used Neotropical army ants (Dorylinae) to investigate the potential earlier existence of land bridges between North and South America. With wingless queens and workers that can only travel on dry ground, army ant colonies are unable to disperse across any larger water bodies (Winston et al. 2016) and are therefore particularly suited to

answer this question. Based on restriction-site associated DNA sequencing (RAD-seq) and a concatenated alignment of 419 804 RAD-seq loci, Winston et al. (2016) generated a time-calibrated phylogeny that supported migration from South to Central America prior to 3 Ma for populations of the four species *Eciton burchellii* (4.3 Ma), *E. vagans* (5.5 Ma), *E. lucanoides* (6.4 Ma), and *E. mexicanum* (6.6 Ma). These estimates appear to support the existence of earlier land bridges, however, the results might be compromised by the fact that concatenation was used for phylogenetic inference. In the presence of incomplete lineage sorting, concatenation has not only been shown to be statistically inconsistent with a tendency to inflate support values (Kubatko and Degnan 2007; Roch and Steel 2014; Linkem et al. 2016), studies based on empirical as well as simulated data have also highlighted that concatenation may lead to branch-length bias and potentially misleading age estimates, particularly for younger divergences (McCormack et al. 2011; Angelis and dos Reis 2015; Mendes and Hahn 2016; Meyer et al. 2016; Ogilvie et al. 2016a,b).

A better alternative for more accurate estimates of divergence times related to the isthmus closure is the multi-species coalescent (MSC) model (Maddison 1997; Ogilvie et al. 2016a,b). While the MSC also does not account for processes like introgression or gene duplication, it incorporates incomplete lineage sorting, which is likely the most prevalent cause of gene-tree heterogeneity in rapidly diverging lineages (Hobolth et al. 2007; Scally et al. 2012; Suh et al. 2015; Edwards et al. 2016; but see Scornavacca and Galtier 2016). Unfortunately, available software implementing the MSC model either does not estimate branch lengths in units of time (Rannala and Yang 2003; Liu 2008; Kubatko et al. 2009; Liu et al. 2010; Bryant et al. 2012; Chifman and Kubatko 2014; Mirarab and Warnow 2015) or is computationally too demanding to be applied to genome-wide data (Heled and Drummond 2010; Ogilvie et al. 2016a). To fill this gap in the available methodology, we here develop a new approach combining the Bayesian species-tree inference of the software SNAPP (Bryant et al. 2012) with a molecular clock model that can be calibrated with

fossil or biogeographic constraints. SNAPP is well suited for analyses of genome-wide data as it infers the species tree directly from single-nucleotide polymorphisms (SNPs), through integration over all possible gene trees on the basis of the MSC model. By using SNPs as markers, SNAPP avoids the issue of within-locus recombination, a common model violation for almost all other implementations of the MSC (Lanier and Knowles 2012; Gatesy and Springer 2013, 2014; Scornavacca and Galtier 2016; Springer and Gatesy 2016; Edwards et al. 2016). SNAPP has been used in over 50 studies (Supplementary Table S1), but with few exceptions, none of these studies inferred absolute divergence times. In three studies that estimated divergence times (Lischer et al. 2014; Demos et al. 2015; Ru et al. 2016), branch lengths were converted to absolute times on the basis of an assumed mutation rate for the SNP set, a practice that should be viewed with caution due to ascertainment bias (Lozier et al. 2016, also see the results of this study). With the possibility to analyze thousands of markers simultaneously, SNAPP nevertheless promises high precision in relative branch-length estimates, and accurate absolute divergence times when properly calibrated with fossil or biogeographic evidence.

We evaluate the accuracy and precision of our approach as well as its computational requirements using an extensive set of simulations, and we compare it to divergence-time estimation based on concatenation. We then apply our method to reanalyze genomic data of Neotropical army ants with the MSC model, and we use it to estimate divergence times of Neotropical sea catfishes (Ariidae) based on newly generated RAD-seq data. Sea catfishes include species endemic to the TEP as well as Caribbean species in several genera. They inhabit coastal brackish and marine habitats down to a depth of around 30 m (Cervigón et al. 1993) and are restricted in dispersal by demersal lifestyle and male mouthbrooding. Sea catfishes are thus directly affected by geographical changes of the coast line which makes them ideally suited to inform about vicariance processes related to the emergence of the Isthmus of Panama.

Bayesian Divergence-Time Estimation with Simulated SNP Data

We designed three experiments based on simulated data to thoroughly test the performance of the MSC model implemented in SNAPP as a tool for divergence-time estimation with SNP data. In experiment 1, we test the accuracy and precision of divergence times estimated with SNAPP and the degree to which these are influenced by the size of the SNP data set and the placement of node-age constraints. We also use the set of analyses conducted for experiment 1 to quantify how computational time requirements of SNAPP analyses depend on these factors. In experiment 2, we further evaluate SNAPP's estimates of divergence times, the molecular clock rate, and the effective population size, based on data sets that include or exclude invariant sites, with or without ascertainment-bias correction. Finally, in experiment 3, we compare divergence-time estimates based on the MSC model implemented in SNAPP with those inferred with concatenated data using BEAST (Bouckaert et al. 2014). Characteristics of all simulated data sets are summarized in Table 1. Based on the results of experiments 1-3, we develop recommendations for divergence-time estimation with SNP data, and we then apply this approach to infer timelines of evolution for Neotropical army ants and sea catfishes.

Table 1: Simulated data sets used in experiments 1-3.

# SNPs	Invariant sites	Ascertainment bias	Calibration	Model (implementation)	Experiment
300	excluded	corrected	root node	MSC (SNAPP)	1
1 000	excluded	corrected	root node	MSC (SNAPP)	1,2,3
3 000	excluded	corrected	root node	MSC (SNAPP)	1
300	excluded	corrected	young node	MSC (SNAPP)	1
1 000	excluded	corrected	young node	MSC (SNAPP)	1,3
3 000	excluded	corrected	young node	MSC (SNAPP)	1
1 000	excluded	not corrected	root node	MSC (SNAPP)	2
1 000	included	not present	root node	MSC (SNAPP)	2
1 000	included	not present	root node	concatenation (BEAST)	3
1 000	included	not present	young node	concatenation (BEAST)	3

Simulating Genome-Wide SNP Data

All simulation parameters, including the number of extant species, the age of the species tree, the population size, the generation time, the mutation rate, and the number of loci per data set were chosen to be roughly similar to those expected in empirical analyses with the software SNAPP (Supplementary Table S1). All simulated data sets were based on the same set of 100 species trees generated with the pure-birth Yule process (Yule 1925) (which is also the only tree prior currently available in SNAPP). Ultrametric species trees conditioned to have 20 extant species were generated with branch lengths in units of generations, using a constant speciation rate $\lambda = 4 \times 10^{-7}$ species/generation. Assuming a generation time of 5 years, this speciation rate translates to $\lambda = 0.08$ species/myr, within the range of speciation rates observed in rapidly radiating vertebrate clades (Alfaro et al. 2009; Rabosky et al. 2013). The ages of the resulting species trees ranged from 2.8 to 12.7 (mean: 6.5) million generations or from 14.2 and 63.6 (mean: 32.3) myr, again assuming the same generation time of 5 years.

For each simulated species tree, 10 000 gene trees with four sampled haploid individuals per species were generated with the Python library DendroPy (Sukumaran and Holder 2010), using a constant effective population size of 50 000 haploid individuals on each branch. Sequences with a length of 200 bp were then simulated along each of the $100 \times 10\,000$ gene trees with the software Seq-Gen (Rambaut and Grassly 1997), according to the Jukes-Cantor model of sequence evolution (Jukes and Cantor 1969) and a rate of 10^{-9} mutations per site per generation or 2×10^{-4} mutations per site per myr. The expected number of mutations per site between two individuals of a panmictic population, Θ, can be calculated as $\Theta = 4N\mu$, where N is the number of diploid individuals, or half the number of haploid individuals, and μ is the mutation rate per site per generation (Felsenstein 1992; Bryant et al. 2012). With the settings used in our simulations ($N = 25\,000$; $\mu = 10^{-9}$), the expected number of mutations per site between two

individuals of the same population is therefore $\Theta = 4 \times 25\,000 \times 10^{-9} = 10^{-4}$.

At least 9 965 (mean: 9 997.9) of the resulting 10 000 alignments per species tree contained one or more variable sites. The mean number of SNPs in the 10 000 alignments per species tree ranged from 5.9 to 15.2 and correlated with the age of the species tree (linear regression; $R^2 = 0.63$, $p < 0.001$). A single SNP was selected at random from all except completely invariable alignments to generate data sets of close to 10 000 unlinked SNPs for each of the 100 species trees. For each species, alleles of the four haploid individuals were combined randomly to form two diploid individuals, which resulted in mean heterozygosities between 0.0012 and 0.0034. The resulting data sets of close to 10 000 unlinked SNPs were further subsampled randomly to generate sets of 300, 1 000, and 3 000 bi-allelic SNPs for each species tree (see Table 1).

For the analyses in experiments 1 and 2, each of the 100 data sets of 300, 1 000, and 3 000 SNPs was translated into the format required for SNAPP, where heterozygous sites are coded with "1" and homozyguous sites are coded as "0" and "2". Per site, the codes "0" and "2" were randomly assigned to one of the two alleles to ensure that the frequencies of these codes were nearly identical in each data set. For experiment 2, in which we tested for the effect of ascertainment bias in SNAPP analyses, the data sets of 1 000 SNPs were also modified by adding invariant sites. To each set of 1 000 SNPs, between 12 184 and 32 740 invariant sites (alternating "0" and "2") were added so that the proportion of SNPs in these data sets matched the mean proportion of variable sites in the alignments initially generated for the respective species tree. Finally, for analyses using concatenation in experiment 3, we added the same numbers of invariant sites to the data sets of 1 000 SNPs, however, in this case we used the untranslated versions of these data sets with the original nucleotide code, and also used nucleotide code for the added invariant sites (randomly selecting "A" , "C", "G", or "T" at each site).

Inferring Divergence Times from Simulated SNP Data

Input data and analysis settings were specified in the XML format of BEAST and SNAPP (Drummond and Bouckaert 2015). We used XML files produced with the SNAPP package for BEAUti as templates, however, several important modifications were made to allow divergence-time estimation with SNAPP. First, the forward and reverse mutation rate parameters u and v were fixed by setting both values to 1.0 and removing all operators that propose changes to these rates. By doing so, we assume a symmetric substitution model as well as equal frequencies, which is justified given that homozygous nucleotide alleles were translated into the codes "0" and "2" at random, independently at each site.

Second, we added a parameter for the rate of a molecular clock, and included a strict clock model, the only clock model currently supported by SNAPP, as a branch-rate model within SNAPP's tree-likelihood XML element. As a single operator on the clock rate, we selected the standard scale operator implemented in SNAPP (described in Drummond et al. 2002). We applied both uniform and one-on-x clock-rate priors in initial tests, but then decided to use exclusively one-on-x clock-rate priors in final analyses, based on preliminary results and the recommendation by one of SNAPP's developers (Remco Bouckaert, priv. comm.). In contrast to a uniform prior, the one-on-x prior gives a preference to smaller rate values and thus helps to avoid unrealistically high clock rate estimates. In addition, the one-on-x prior, just like uniform priors, has the advantage of being scale-independent, which means that it can be applied equally in analyses of fast or slow-evolving groups. The molecular clock rate was calibrated through age constraints on a single node of the species tree. To compare the effects of old and young calibrations, we conducted separate sets of analyses in which we placed this age constraint either on the root node or on the node with an age closest to one third of the root age (see Table 1). In each case, calibration nodes were constrained with log-normal calibration densities centered on the true node age. Specifically, these calibration densities were parameterized with an

offset of half the true node age, a mean in real space of half the true node age, and a standard deviation of the log-transformed distribution of 0.1. These node-age constraints were specified in the XML input file for SNAPP exactly as they would be in any analysis with BEAST 2.

Third, initial tests indicated that our simulated data sets contained very little information about the ancestral population sizes on internal branches of the species tree, and that unreliable estimation of these population sizes could confound divergence-time estimates. We therefore decided not to estimate the population-size parameter Θ individually for each branch as is usually done in SNAPP analyses, but instead to estimate just a single value of Θ for all branches, assuming equal population sizes in all species. This assumption was met in our simulated data sets but may often be violated by empirical data sets; we turn to the implications of this violation in the Discussion. While SNAPP does not explicitly provide an option to link the Θ values of all branches, we achieved this linking by replacing all operators on Θ with a single scale operator that changes all Θ values by the exact same scale value whenever its propsals are accepted. As a prior on Θ, four distribution types are available in SNAPP: uniform, gamma, inverse gamma, and the auto-correlated "CIR" process (Cox et al. 1985; Lepage et al. 2007). While gamma-distributed priors on Θ are most commonly used in SNAPP analyses, we selected the less informative and scale-independent uniform prior for our analyses. The boundaries of SNAPP's uniform prior on Θ are indirectly specified in the program's source code as limits to the coalescent rate $2/\Theta$, which are defined as $0 \leq 2/\Theta \leq 10^9$ and thus constrain Θ to $\Theta \geq 2 \times 10^{-9}$. Currently, this boundary can not be modified by the user, however, none of the estimated Θ values in our analyses came close to this limit, therefore we assume that SNAPP's uniform prior on Θ is entirely uninformative in practice. Through experimental modification of the source code of SNAPP, we initially also implemented and tested a one-on-x prior on Θ in preliminary analyses. However, as this prior applies independently

to the linked Θ values of all species-tree branches (a total of 38 branches in our simulated species trees), its effect appeared too strong and led to severe parameter underestimation (each division of Θ in half multiplied the overall prior probability by 2^{38}, equivalent to an increase by 11.4 log units). Thus, we considered the uniform prior on Θ to be appropriate when the number of species-tree branches is large relative to the number of individuals sampled from each species, as was the case in all our analyses.

Finally, as we were interested in SNAPP's ability to infer divergence times rather than the species-tree topology (which has been demonstrated previously; Bryant et al. 2012), we fixed the species-tree topology by using the true species tree as a starting tree and by removing the "NodeSwapper" operator, the only operator on the species-tree topology. We provide a script written in Ruby, "snapp_prep.rb", to generate XML input files for SNAPP corresponding to the settings described above (with or without a fixed species tree), given a nucleotide data set in phylip format, a table assigning individuals to species, one or more node-age constraints, and (optionally) a starting tree. This script is freely available at https://github.com/mmatschiner/snapp_prep.

As SNAPP is specifically designed for the analysis of bi-allelic SNPs, its algorithm explicitly accounts for ascertainment bias introduced by the exclusion of invariable sites (Bryant et al. 2012; RoyChoudhury and Thompson 2012). Nevertheless, SNAPP allows invariant sites in the data set and the user may specify whether or not these have been excluded by setting the option "non-polymorphic" to "true" (invariant sites are included) or "false" (invariant sites are excluded). Accordingly, we set this option to "false" for all analyses of experiment 1, but not for the analyses of experiment 2 in which either ascertainment bias was not corrected for or invariant sites were added to data sets of 1 000 SNPs (see Table 1). This option did not apply to the analyses of concatenated data in experiment 3 as these were not conducted with SNAPP. As a substitution model, we applied the HKY model (Hasegawa et al. 1985) in analyses of concatenated data.

All XML files were analyzed using BEAST v.2.3.0 (Bouckaert et al. 2014) either with the SNAPP package v.1.3.0 (all analyses of experiment 1 and 2) or without additional packages (analyses of concatenated data sets in experiment 3). Analyses using SNAPP were conducted on the Abel computing cluster provided by the University of Oslo, using four threads on dual eight-core Intel Xeon E5-2670 (Sandy Bridge-EP) CPUs running at 2.6 GHz. We performed 1-9 million Markov-chain Monte Carlo (MCMC) iterations per SNAPP analysis (1 million for analyses with root node calibrations, 2-9 million for analyses with calibrations on younger nodes) and 500 000 iterations per concatenation analysis. Stationarity of MCMC chains was assessed by calculating effective samples sizes (ESS) for all parameters with the R package CODA v.0.17-1 (Plummer et al. 2006), after discarding the first 10% of the chain as burn-in. For each SNAPP analysis, we recorded i) the time required per MCMC iteration, ii) the number of iterations required for convergence (ESS $\geq$ 200) of all parameters, and iii) the time required for convergence as the product of i) and ii).

Results: Precision and Accuracy of Parameter Estimates Based on Simulated SNPs

Experiment 1.— A comparison of true and estimated node ages, for analyses of 100 data sets of 300, 1 000, and 3 000 SNPs with node-age constraints on either the root or a younger node, is shown in Figure 1 and summarized in Table 2. As measured by the width of 95% highest posterior density (HPD) intervals, precision was generally greater for younger nodes and increased when larger numbers of SNPs were used for the analysis. In all sets of analyses, over 95% of the 95% HPD intervals contained the true age of the node, indicative of accurate inference free of node-age bias (Heath et al. 2014; Gavryushkina et al. 2014; Matschiner et al. 2016). The percentage of 95% HPD intervals containing the

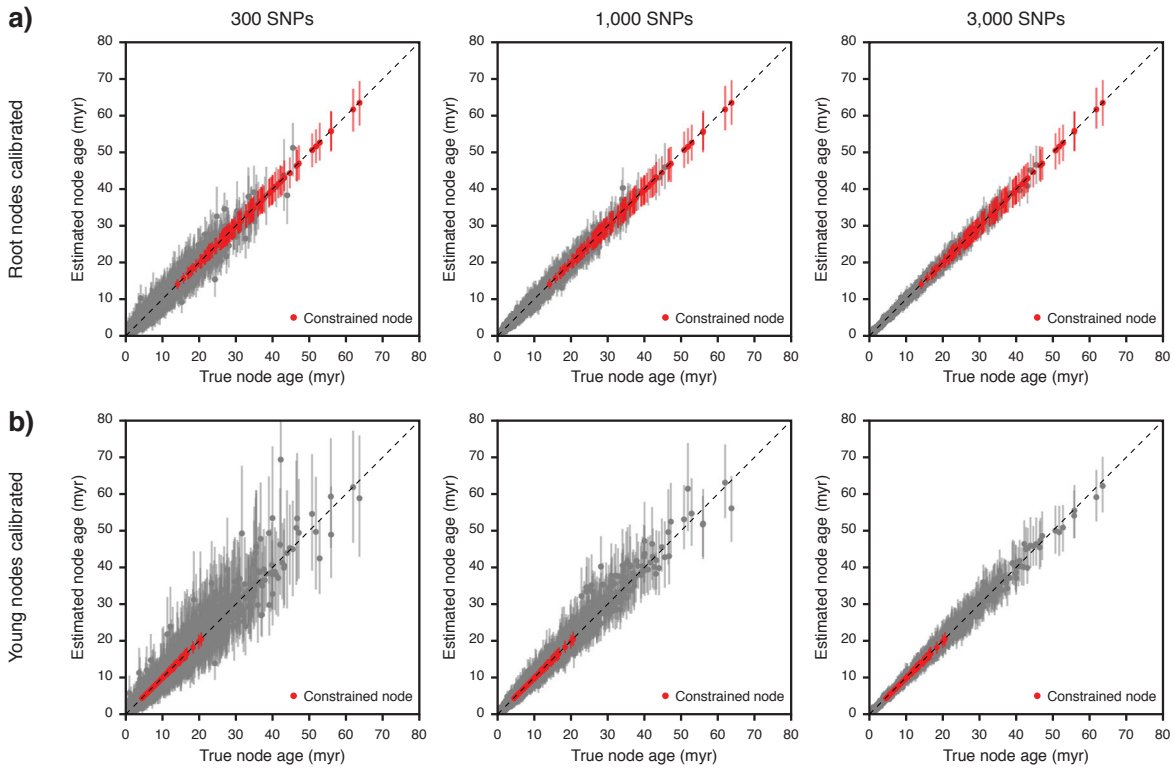


Figure 1: Comparison of true and estimated node ages (experiment 1).
Results are based on 100 species trees and 300 to 3 000 SNPs generated per species tree. a) Node ages estimated with an age constraint on the root. b) Node ages estimated with an age constraint on a node that is approximately a third as old as the root. Mean age estimates of constrained and unconstrained nodes are marked with red and gray circles, respectively, and vertical bars indicate 95% HPD intervals.

true age was always slightly higher in analyses with root-node constraints even though the width of these HPD intervals was generally smaller.

The computational time required per MCMC iteration depended less on the number of SNPs (linear regression without intercept; $R^2 = 0.83$, $p < 0.001$) than on the number of unique site patterns in each data set ($R^2 = 0.96$, $p < 0.001$). The data sets of 300, 1 000, and 3 000 SNPs included 64 to 100 (mean: 81.8), 112 to 166 (mean: 138.0), and 183 to 280 (mean: 232.4) unique site patterns, respectively, and therefore showed a less-than-linear increase of the number of unique sites with the number of SNPs (Figure 2). The number of

Table 2: Accuracy and precision of node-age estimates (experiment 1).

Mean 95% HPD width (myr):

		True node age (myr)			
# SNPs	Calibration	0-20	20-40	40-60	All
300	root node	4.85	8.62	9.93	5.46
1 000	root node	2.98	6.64	9.29	3.59
3 000	root node	2.11	5.71	8.97	2.71
300	young node	5.57	16.47	26.74	7.40
1 000	young node	3.43	10.74	16.75	4.65
3 000	young node	2.31	7.44	11.99	3.17

Percentage of 95% HPD containing the true node age:

		True node age (myr)			
# SNPs	Calibration	0-20	20-40	40-60	All
300	root node	95.8	97.4	100.0	96.1
1 000	root node	96.6	99.3	100.0	97.1
3 000	root node	97.3	100.0	100.0	97.7
300	young node	95.8	96.4	95.2	95.9
1 000	young node	96.7	95.6	100.0	96.6
3 000	young node	97.1	99.3	100.0	97.5

MCMC iterations required for convergence (all ESS values ≥ 200) increased with the number of SNPs when constraints were placed on the root node, but decreased when younger age constraints were used (Table 3). As a result, the computational time required for convergence was driven by a combination of both the size of the data set and the placement of the node-age constraint. On average, convergence was reached after 11.7 hours when the data set contained 300 SNPs and the phylogeny was time-calibrated with an age constraint on the root node, but it required on average 94.0 hours for convergence when the same number of SNPs were analyzed with an age constraint on a younger node. In contrast, time to convergence was less variable with larger data sets of 3 000 SNPs, which required on average 46.3 hours with root-node constraints and 70.2 hours with constraints on younger nodes.

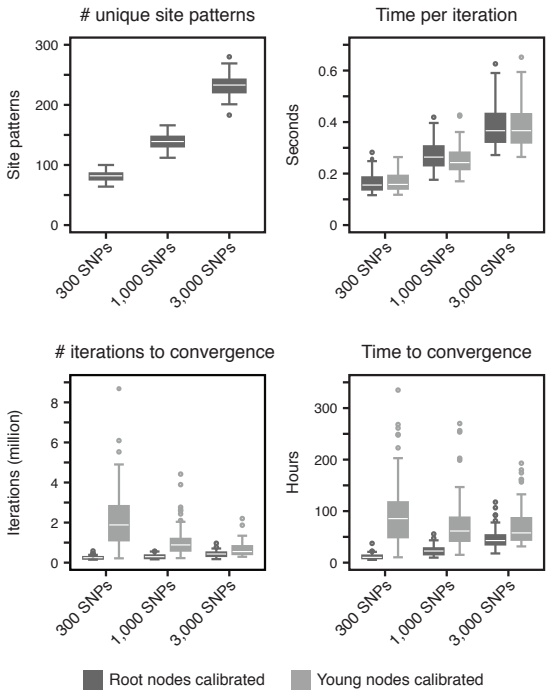


Figure 2: Comparison of run times of SNAPP analyses (experiment 1).
The number of unique site patterns, the time required per MCMC iteration, the number of iterations required for convergence (ESS values for all parameters ≥ 200), and the resulting time required for convergence are compared for analyses of data sets comprising 300, 1 000, and 3 000 SNPs.

Table 3: Mean computational time requirements of SNAPP analyses (experiment 1).

# SNPs	Calibration	Time per iteration	# iterations to convergence	Time to convergence
300	root node	0.16 secs	254 175	11.7 hrs
1 000	root node	0.27 secs	309 455	23.2 hrs
3 000	root node	0.38 secs	433 035	46.3 hrs
300	young node	0.17 secs	2 056 755	94.0 hrs
1 000	young node	0.25 secs	1 030 520	72.4 hrs
3 000	young node	0.38 secs	662 430	70.2 hrs

Experiment 2.— The comparison of analyses with and without SNAPP's ascertainment-bias correction showed that this correction is required for accurate node-age estimates when invariant sites are excluded. Without ascertainment-bias correction, only 86.6% of the 95% HPD intervals contained the true node age (Figure 3a) (compared to

97.1% when ascertainment bias is corrected; Table 2), while the width of these intervals remained similar (3.54 compared to 3.59 myr). Of the 13.4% of 95% HPD intervals that did not contain the true node age, almost all (13.2%) were younger than the true node age, indicating a tendency to underestimate node ages when ascertainment bias is not taken into account.

Regardless of whether SNAPP's ascertainment-bias correction was used or not, the clock rates and Θ values estimated from data sets without invariant sites did not match the settings used for simulations (clock rate $= 2 \times 10^{-4}$ mutations per site per myr; $\Theta = 10^{-4}$; see above) (Fig. 3b,c, Table 4). While both parameters were underestimated roughly by a factor of three when ascertainment bias was corrected for, leaving this bias unaccounted led to parameter overestimation by more than an order of magnitude. Importantly, however, when ascertainment bias was accounted for, the resulting estimates of the population size N (calculated as $N = \Theta/4\mu$ with μ being the mutation rate per generation, i.e., the estimated clock rate divided by the number of generations per myr) accurately recovered the true population size used for simulations ($N = 25\,000$; see above), as exactly 95% of the 95% HPD intervals included the true parameter value (Fig. 3d, Table 4). In contrast, the population size was underestimated when ascertainment bias was not corrected for: mean estimates were on average 17.4% lower than the true population size and 35% of the 95% HPD intervals did not include the true parameter value (Fig. 3d, Table 4).

Table 4: Mean estimates of clock rate, Θ, and the population size, in analyses of data sets with and without ascertainment bias (experiment 2).

Invariant sites	Ascertainment bias	Clock rate (myr^{-1})	Θ	Population size
excluded	corrected	5.93×10^{-5}	2.89×10^{-5}	24 438
excluded	not corrected	5.02×10^{-3}	2.05×10^{-3}	20 644
included	not present	1.99×10^{-4}	9.59×10^{-5}	24 226

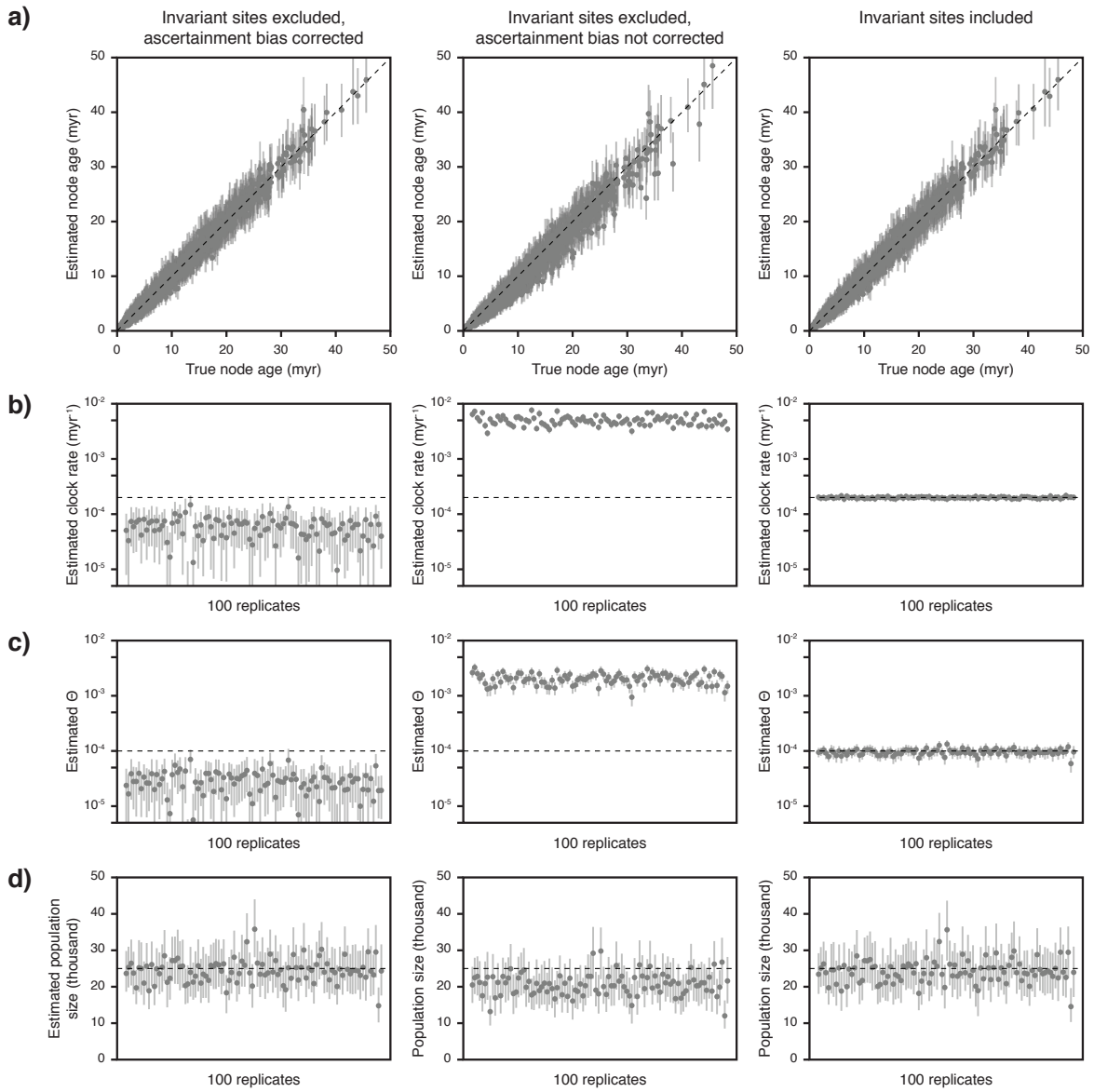


Figure 3: Estimates of node ages, the clock rate, Θ, and the population size, with and without ascertainment bias (experiment 2).
Results are based on data sets of 1 000 SNPs generated for each of 100 species trees, analyzed with and without SNAPP's ascertainment-bias correction or after adding invariant sites to the data sets. Gray circles indicate mean estimates and 95% HPD intervals are marked with vertical bars. The visualization of node-age estimates in a) is equivalent to the illustration in Fig. 1, except that only unconstrained nodes are shown.

Instead of accounting for ascertainment bias, the addition of invariant sites also allowed the accurate estimation of divergence times and population sizes, with 97.2% and 94% of the true node ages and the true population sizes contained within 95% HPD intervals, respectively (Fig. 3a,d). Moreover, the true clock rate and the true Θ value were also recovered reliably in these analyses and were included in 100% and 90% of the 95% HPD intervals (Fig. 3b,c, Table 4).

Experiment 3.— The comparison of node-age errors resulting from analyses with the MSC and with concatenation indicated that both methods perform equally well for nodes older than a few million years. However, the relative error in estimates of young nodes ($\leq$ 3 myr) was markedly greater with the concatenation approach (Fig. 4), regardless of whether species trees were time calibrated with age constraints on root nodes or young nodes (Table 5). In concatenation analyses with age constraints on root nodes, the ages of nodes younger than 10 myr were on average misestimated by 48.9%, compared to 17.3% in analyses with the MSC. Similarly, when age constraints were placed on young nodes, concatenation led to an average error of 58.5% for nodes younger than 10 myr, while analyses with the MSC resulted in misestimation by 20.8% for the same nodes.

Table 5: Mean error in node-age estimates in analyses using the MSC or concatenation, given in percent deviation from the true node age (experiment 3).

		True node age (myr)					
Calibration	Model (implementation)	0-10	10-20	20-30	30-40	40-50	All
root node	MSC (SNAPP)	17.3	5.8	4.2	3.6	1.5	12.8
root node	concatenation (BEAST)	48.9	5.8	4.2	3.6	1.5	32.5
young node	MSC (SNAPP)	20.8	7.5	7.4	6.7	5.6	15.4
young node	concatenation (BEAST)	58.5	7.1	7.0	6.9	6.0	37.9

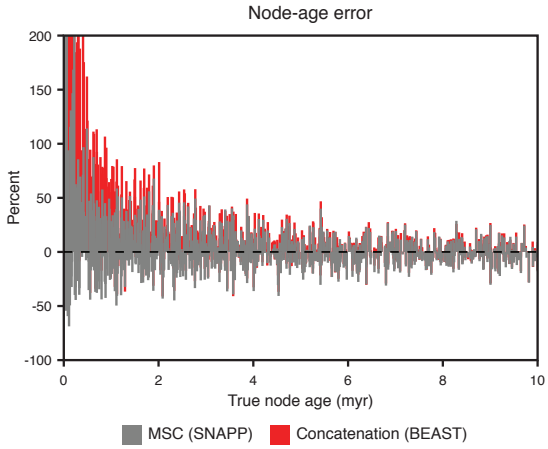


Figure 4: Error in node-age estimates obtained with the MSC or with concatenation (experiment 3).
Results are based on analyses of 100 data sets of 1 000 SNPs, with age constraints on the root node of each species tree (qualitatively similar results were obtained with age constraints on younger nodes; not shown). Vertical bars visualize the degree to which node ages are overestimated in percent of the true node age. Gray and red bars indicate node age estimates obtained with the MSC implemented in SNAPP and with BEAST analyses of concatenated data, respectively. Only nodes with true ages up to 10 myr are shown to highlight differences between the two methods.

Bayesian Divergence-Time Estimation with Empirical SNP Data

Reanalysis of Neotropical Army Ant SNP Data

Divergence times of Neotropical army ants were estimated by Winston et al. (2016) based on a data set of 419 804 RAD-seq loci (39 927 958 bp with 87.2% missing data), sequenced from 146 specimens of 18 species in five genera. Phylogenetic analysis of the concatenated data set led to divergence-time estimates older than 3 Ma between Central American and predominantly South American populations in each of four species of genus *Eciton* (*E. mexicanum*, *E. lucanoides*, *E. vagans*, and *E. burchellii*), which were taken as evidence for temporary land bridges prior to the full closure of the Panamanian Isthmus

(Winston et al. 2016). To allow an efficient reanalysis of army ant divergence times with the MSC model, we reduced the size of this data set to the four specimens with the lowest proportions of missing data for each species, or for each of the two geographic groups in the four species *E. mexicanum*, *E. lucanoides*, *E. vagans*, and *E. burchellii*. We further filtered the data set so that maximally one SNP was included per RAD locus. The reduced data set included 413 bi-allelic SNPs suitable for analysis with SNAPP, with data available for at least one specimen per species. SNAPP input files in XML format were generated with the script "snapp_prep.rb" (see above), using the same settings as for analyses of simulated data, except that the operator on the tree topology was not excluded. As in Winston et al. (2016), time calibration was based on the published age estimate of 37.23 Ma (confidence interval: 46.04-28.04 Ma) for the most recent common ancestor of Neotropical army ants (Brady et al. 2014). We specified this age constraint as a normally-distributed calibration density with a mean of 37.23 Ma and a standard deviation of 4.60 myr. To further reduce computational demands of the SNAPP analysis, we also enforced monophyly of each genus, and of each of the four species represented by two populations, according to the strong support (BPP: 1.0) that these groups received in Winston et al. (2016). We performed five replicate SNAPP analyses, each with a run length of 500 000 MCMC iterations. Chain convergence and stationarity were assessed through comparison of parameter traces among analysis replicates, using the software Tracer v.1.6 (Rambaut et al. 2014). As stationarity was supported by ESS values above 200 for all parameters in each analysis, MCMC chains of analysis replicates were combined after discarding the first 10% of each chain as burn-in. None of the ESS values of the combined chains were below 1000, strongly supporting convergence of all analyses.

For comparison, we also repeated the analysis of army ant divergence times based on concatenation of all sequences, using a single specimen for each of the 22 species and excluding alignment positions with more than 50% missing data, which resulted in an

alignment of 3 058 724 bp (with 37.1% missing data). Analyses based on concatenation were conducted in BEAST, using the GTR substitution model (Tavaré 1986) with gamma-distributed among-site rate variation and the same tree prior, clock model, and constraints as in analyses with the MSC. We again performed five analysis replicates, each with 600 000 MCMC iterations, and stationarity and convergence were again supported by ESS values above 200 in each individual analysis replicate and above 1 000 after combining the five MCMC chains.

Results: Timeline of Neotropical Army Ant Diversification

Our reanalysis of Neotropical army ant SNP data with the MSC resulted in a strongly supported phylogeny (mean BPP: 0.94) that recovered the topology proposed by Winston et al. (2016) with the single exception that *Eciton mexicanum* appeared as the sister of *E. lucanoides* rather than diverging from the common ancestor of *E. lucanoides*, *E. burchellii*, *E. drepanophorum*, and *E. hamatum* (Supplementary Figure S1 and Supplementary Table S2). However, the timeline of army ant divergences inferred with the MSC was markedly different from the timeline estimated by Winston et al. (2016). Whereas Winston et al. (2016) estimated the crown divergence of the genus *Eciton* to have occurred around 14.1 Ma, our analysis based on the MSC placed this divergence around the Miocene-Pliocene boundary (5.48 Ma; 95% HPD: 7.52-3.52 Ma). In contrast to the previous analysis, the divergences between Central American and predominantly South American populations within *E. mexicanum* (1.82 Ma; 95% HPD: 3.02-0.76 Ma), *E. lucanoides* (2.47 Ma; 95% HPD: 3.88-1.22 Ma), *E. vagans* (0.33 Ma; 95% HPD: 0.71-0.05 Ma), and *E. burchellii* (0.54 Ma; 95% HPD: 1.12-0.13 Ma) were all placed in the Pleistocene in our study, in agreement with migration subsequent to the final isthmus closure. The population size inferred with the MSC, applying to all extant and ancestral

species equally, was 53 854 (95% HPD: 34 433-75 294), based on an assumed generation time of 3 years (Berghoff et al. 2008).

When using concatenation to estimate army ant divergence times, the mean age estimates of splits between Central American and predominantly South American lineages within *E. mexicanum* (2.47 Ma; 95% HPD: 3.09-1.88 Ma), *E. lucanoides* (3.74 Ma; 95% HPD: 4.68-2.83 Ma), *E. vagans* (1.31 Ma; 95% HPD: 1.65-1.00 Ma), and *E. burchellii* (2.07 Ma; 95% HPD: 2.56-1.55 Ma) were 35.7-397.1% older (Supplementary Figure S2 and Supplementary Table S3) than those based on the MSC model. While these age estimates for population splits in *E. mexicanum*, *E. vagans*, and *E. burchellii* would still agree with migration after the final closure of the isthmus, the confidence interval for the divergence time of populations within *E. lucanoides* does not include the accepted age for the final isthmus closure (2.8 Ma; O'Dea et al. 2016) and would thus support the existence of earlier land bridges.

Generation of SNP Data for Neotropical Sea Catfishes

Twenty-six individuals that belong to 21 recognized species and two possibly cryptic species of the five Neotropical sea catfish genera *Ariopsis*, *Bagre*, *Cathorops*, *Notarius*, and *Sciades* were analyzed using RAD-seq (samples listed in Table 6, including GPS coordinates and locality names). For four of these genera, our taxon set includes both species endemic to the TEP and species endemic to the Caribbean, hence, the divergences of these taxa were expected to have occurred prior to or simultaneously with the closure of the Panamanian Isthmus. Taxonomic identifications have previously been conducted for the same samples based on morphology as well as mitochondrial sequences (see Stange et al. 2016 for details) and were therefore considered to be reliable.

Fresh fin tissues were preserved in 96% ethanol for subsequent DNA extraction. DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen, Valencia, USA)

following the manufacturer's instructions. RNase treatment after digestion (but before precipitation) was performed in order to improve the purity of the samples. DNA concentrations were measured using a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, Waltham, MA, USA). The samples were standardized to 23.5 ng/μl and used to generate a RAD library, following the preparation steps described in Roesti et al. (2012) and using restriction enzyme Sbf1. We assumed a genome size of approximately 2.4 Gb as inferred from available C-values for sea catfishes (Gregory 2016). Therefore, we expected a recognition site frequency of 20 per Mb, which would yield around 50,000 restriction sites in total. Specimens were individually barcoded with 5-mer barcodes.

Two libraries were prepared and single-end sequenced with 201 cycles on the Illumina HiSeq 2500 platform, at the Department of Biosystems Science and Engineering, ETH Zurich. The resulting raw reads were demultiplexed (NCBI study accession: SRP086652) based on the individual barcodes with the script "process_radtags.pl" of the software Stacks v.1.32 (Catchen et al. 2011) and further analyzed with pyRAD v.3.0.5 (Eaton 2014). Settings of the pyRAD analysis included a minimum depth of 20 per within-sample cluster (Mindepth: 20), a maximum of four sites with a quality value below 20 (NQual: 4), maximally 20 variable sites within a cluster (Wclust: 0.89), and a minimum of 18 samples in a final locus (MinCov: 18). Quality filtering (step 2 in the pyRAD pipeline; Eaton 2014) resulted in the exclusion of 23-56% of the reads; after filtering, between 2.4 and 5.6 million reads remained per individual. Reads that passed the applied filtering steps resulted in about 40 000 to 166 000 within-sample clusters (step 3) with mean depths between 44 and 89. The estimated error rate and heterozygosity of these clusters (step 4) amounted to 0.0004-0.0009 and 0.0042-0.0107, respectively. Consensus-sequence creation from the within-sample clusters (step 5) based on the estimated heterozygosity and error rate, with a maximum of 20 variable sites, a minimal depth of 20, and additional paralog filtering (maximally 10% shared heterozygous sites), resulted in 21 575 to 38 182

Table 6: Taxa sampled for phylogenetic inference. Localities and sampling coordinates are provided for sampling sites in Panama and Venezuela.

Scientific name	Locality	Sampling coordinates		Specimen ID
Ariopsis sp. nov. (*A. jimenezi*)	Pearl islands / Casaya island, Panama, PA	8°34′38.6″N	79°03′03.6″W	08E4
Ariopsis seemanni	Rio Estero Salado, Cocle, PA	8°10′30.3″N	80°29′35.1″W	08C8
Bagre bagre	Lago de Maracaibo / Isla de San Carlos, Zulia, VE	10°59′55.1″N	71°36′19.8″W	05A9
Bagre marinus	Gulf of Venezuela, Falcon, VE	11°14′15.3″N	70°30′53.1″W	05I5
Bagre panamensis	Rio Estero Salado, Cocle, PA	8°10′30.3″N	80°29′35.1″W	08C1
Bagre pinnimaculatus 1	Rio Estero Salado, Cocle, PA	8°10′30.3″N	80°29′35.1″W	08B7
Bagre pinnimaculatus 2	Gulf of Panama	8°48′56.6″N	79°22′50.9″W	01B6
Cathorops fuerthii	Rio Parita, Herrera, PA	8°01′13.7″N	80°27′11.2″W	08D5
Cathorops hypophthalmus	Rio Estero Salado, Cocle, PA	8°10′30.3″N	80°29′35.1″W	08C4
Cathorops nuchalis	Lago de Maracaibo / Puerto Concha, Zulia, VE	9°05′46.0″N	71°42′52.0″W	01G1
Cathorops tuyra	Puente del Rio Chagres,Colon, PA	9°11′34.7″N	79°39′09.4″W	08G2
Cathorops wayuu	Lago de Maracaibo / Guarico, Zulia, VE	10°43′52.0″N	71°31′40.2″W	05C7
Cathorops wayuu	Lago de Maracaibo / Isla de Toas, Zulia, VE	10°57′09.5″N	71°38′49.5″W	05G8
Notarius biffi	Rio San Pedro, Veraguas, PA	7°50′59.2″N	81°07′04.0″W	08A9
Notarius cookei	Rio Santa Maria, Cocle, PA	8°06′20.3″N	80°33′16.1″W	08G9
Notarius grandicassis	Gulf of Venezuela, Falcon, VE	11°14′15.3″N	70°30′53.1″W	05H9
Notarius kessleri	Rio San Pedro, Veraguas, PA	7°50′59.2″N	81°07′04.0″W	08A7
Notarius planiceps	Rio Estero Salado, Cocle, PA	8°10′30.3″N	80°29′35.1″W	08C2
Notarius quadriscutis	Clarines, Anzoategui, VE	10°03′46.8″N	65°11′05.2″W	03F1
Notarius dowii	Puerto Caimito, Panama, PA	8°52′18.9″N	79°42′33.0″W	08D9
Notarius dowii	Rio Santa Maria, Cocle, PA	8°06′20.3″N	80°33′16.1″W	08H1
Sciades herzbergii 1	Clarines, Anzoategui, VE	10°03′46.8″N	65°11′05.2″W	03C6
Sciades herzbergii 2	Gulf of Venezuela, Falcon, VE	11°14′15.3″N	70°30′53.1″W	04D9
Sciades parkeri	Ciudad Bolivar, Bolívar, VE	8°08′51.5″N	63°32′10.7″W	06B9
Sciades proops	Gulf of Venezuela, Falcon, VE	11°14′15.3″N	70°30′53.1″W	05I8
Sciades proops	Puerto La Cruz, Anzoategui, VE	10°12′58.6″N	64°38′39.2″W	06A9

PA, Panama; VE, Venezuela. Specimens of nominal species that are considered independent lineages in our phylogenetic analysis are marked with numbers 1,2.

consensus loci per sample. Between-sample clusters (step 6) were created with the same settings as within-sample clusters. These clusters were filtered again for potential paralogs (step 7) with a maximum of five shared heterozygous sites. The final data set contained 10 991-14 064 clusters per individual. From these clusters, one SNP per locus was selected at random for use in phylogenetic inference, assuming that SNPs of different loci are effectively unlinked.

Inferring Divergence-Times of Neotropical Sea Catfishes

To incorporate existing estimates of the timeline of Neotropical sea catfish evolution into our analyses, we identified the age of the most recent common ancestor of the five sea catfish genera included in our taxon set (*Ariopsis*, *Bagre*, *Cathorops*, *Notarius*, and *Sciades*) from the time-calibrated phylogeny of Betancur-R. et al. (2012). Details of this phylogenetic analysis are given in Betancur-R. et al. (2012). In brief, Betancur-R. et al. (2012) used concatenation of five mitochondrial and three nuclear genes (a total of 7 190 sites) for phylogenetic inference of 144 species (representing 28 of the 29 valid genera of sea catfishes as well as teleost outgroups), and divergence times were estimated with BEAST v.1.6.1 (Drummond et al. 2012) on the basis of 14 teleost fossils and five biogeographic node-age constraints. However, as three of these biogeographic constraints were derived from an assumed closure of the Isthmus of Panama between 3.1 and 2.8 Ma and since our goal was to compare the timeline of Neotropical sea catfish evolution with the age estimates for the closure of the isthmus, we repeated the analysis of Betancur-R. et al. (2012) excluding these three constraints to avoid circular inference. All other analysis settings were identical to those used in Betancur-R. et al. (2012) but we used BEAST v.1.8.3, the latest version of BEAST compatible with the input file of Betancur-R. et al. (2012), and 150 million MCMC iterations for the inference.

The resulting age estimate for the most recent common ancestor of the genera *Ariopsis*, *Bagre*, *Cathorops*, *Notarius*, and *Sciades* (27.42 Ma; 95% HPD: 30.89-24.07 Ma) was then used as a constraint on the root of a species tree of Neotropical sea catfishes inferred with SNAPP, based on our RAD-seq data set of 21 sea catfish species. For this analysis, we used 1 768 bi-allelic SNPs for which data were available for at least one individual of each species or population. *Bagre pinnimaculatus* from Panama and *Sciades herzbergii* from Venezuela were represented by two individuals each, which were both considered as representatives of separate lineages in the SNAPP analyses. Differentiation between the populations from which these individuals were sampled was previously described based on morphology (*Bagre pinnimaculatus*) and distinct mitochondrial haplotypes (both species) (Stange et al. 2016). We again used our script "snapp_prep.rb" (see above) to convert the SNP data set into SNAPP's XML format.

The strict molecular clock rate was calibrated with a normally distributed calibration density (mean: 27.4182 Ma, standard deviation: 1.7 myr) on the root age, according to our reanalysis of the Betancur-R. et al. (2012) data set. In addition, the fossil record of sea catfishes was used to define minimum ages for several lineages. The oldest fossil records of the genera *Bagre*, *Cathorops*, and *Notarius* have been described from the eastern Amazon Pirabas Formation on the basis of otolith and skull material (Aguilera et al. 2013). As the Pirabas Formation is of Aquitanian age (Aguilera et al. 2013), we constrained the divergences of each of the three genera with a minimum age of 20.4 Ma (Cohen et al. 2013). Furthermore, skull remains of the extant species *Sciades dowii*, *Sciades herzbergii*, *Bagre marinus*, and *Notarius quadriscutis* have been identified in the Late Miocene Urumaco Formation of northwestern Venezuela (Aguilera and de Aguilera 2004a), which therefore provides a minimum age of 5.3 Ma for these species. All fossils used for phylogenetic analyses are summarized in Table 7.

We carried out five replicate SNAPP analyses, each with a run length of one million

MCMC iterations, of which the first 10% were discarded as burn-in. Convergence was suggested by ESS values for all parameters above 200 in individual replicate analyses, and by ESS values above 1 000 after combining the output of the five replicates. The combined analysis output was used to sample a set of 1 000 trees as representative of the posterior tree distribution.

Stochastic Mapping of Geographic Distributions

For the purpose of reconstructing ancestral distributions of sea catfishes taking into account the localities of fossil finds, eight fossil taxa were added to each of the 1 000 trees of the posterior tree set, according to their taxonomic assignment: *Cathorops goeldii* (Aguilera et al. 2013) was attached to the stem branch of the genus *Cathorops* with a tip age of 20.4 Ma, reflecting the minimum age of the Pirabas Formation from which this fossil was described. Similarly, *Notarius* sp. and *Bagre protocaribbeanus* from the same formation (Aguilera et al. 2013) were added to the stem branches of the genera *Notarius* and *Bagre*, respectively. In addition, *Bagre protocaribbeanus* from the Venezuelan Cantaure Formation was added as the sister of the older *B. protocaribbeanus* fossil from the Pirabas Formation, with a tip age of 16.0 Ma. Finally, fossil representatives of the species *Sciades dowii*, *Sciades herzbergii*, *Bagre marinus*, and *Notarius quadriscutis* were added as sister branches of these extant taxa, with a tip age corresponding to the minimum age of the Urumaco Formation, 5.3 Ma (see Table 7). For all additions, the age of the attachment points were chosen at random between the fossil's age and the age of the branch to which the fossil was attached. The posterior tree set including fossil attachments was summarized as a maximum-clade-credibility (MCC) tree with node heights according to mean ages (Heled and Bouckaert 2013) (Supplementary Figure S3 and Supplementary Table S4).

Table 7: Fossils of Neotropical sea catfishes used in phylogenetic analyses.

Species	Type	Formation	Locality	Epoch	Age (Ma)
Cathorops goeldii[1]	Skull	Pirabas Fm.	Atalaia Beach, Salinpolis Municipality, Para State, BR	Aquitanian[3]	23.0-20.4[4]
Notarius sp.[2]	Otolith				
Bagre protocaribbeanus[2]	Otolith				
Bagre protocaribbeanus[1]	Otolith	Cantaure Fm.	San José de Cocodite, VE	Burdigalian[5]	20.4-16.0[4]
Sciades dowii[6]	Skull	Urumaco Fm.	Urumaco, VE	Late Miocene[7]	11.6-5.3[4]
Sciades herzbergii[6]	Skull				
Bagre marinus[6]	Skull				
Notarius quadriscutis[6,8]	Skull				

BR, Brazil; VE, Venezuela.
[1]Aguilera et al. (2013); [2]Aguilera et al. (2014); [3]Aguilera et al. (2016); [4]Cohen et al. (2013); [5]Carrillo-Briceño et al. (2016); [6]Aguilera and de Aguilera (2004a); [7]Aguilera and Marceniuk (2012); [8]Originally described as *Aspistor quadriscutis*, however, genus *Aspistor* is here considered synonymous with *Notarius*, following Betancur-R. and Acero P. (2004).

We then used the posterior tree set including fossil taxa to infer the ancestral distribution of sea catfish lineages in the TEP or the Caribbean, based on stochastical mapping of discrete characters (Huelsenbeck et al. 2003) as implemented in function "make.simmap" of the phytools R package (Revell 2012). For this analysis, we assumed a uniform prior probability for the state of the root node (pi="equal") and used an empirically determined rate matrix (Q ="empirical").

Results: Timeline of Neotropical Sea Catfish Diversification

The posterior distribution of species trees is illustrated in Figure 5 in the form of a cloudogram (Bouckaert 2010) with branches colored according to the stochastic mapping of geographic distribution. Our results suggest that the genus *Cathorops* is the outgroup to the other four genera (Bayesian posterior probability, BPP: 1.0) and that the earliest divergence between these groups probably occurred in what is now the Caribbean (BPP: 0.81). The four genera *Notarius*, *Bagre*, *Sciades*, and *Ariopsis* diverged (probably in this order; BPP: 0.92) in a series of rapid splitting events that occurred between 22 and 19 Ma, most likely also in the Caribbean (BPP: 0.89-1.0). Within-genus diversification of the sampled extant lineages began between 12 (*Notarius*) and 5 (*Ariopsis*) Ma, and these initial within-genus divergences occurred both within the Caribbean (*Sciades*, BPP: 1.0; *Bagre*, BPP: 0.77) and the TEP (*Ariopsis*, BPP: 0.89; *Cathorops*, BPP: 0.80). Assuming a generation time of 2 years for sea catfishes (Betancur-R. et al. 2008; Meunier 2012), the estimated population size was 127 250 (95% HPD: 105 120-151 900).

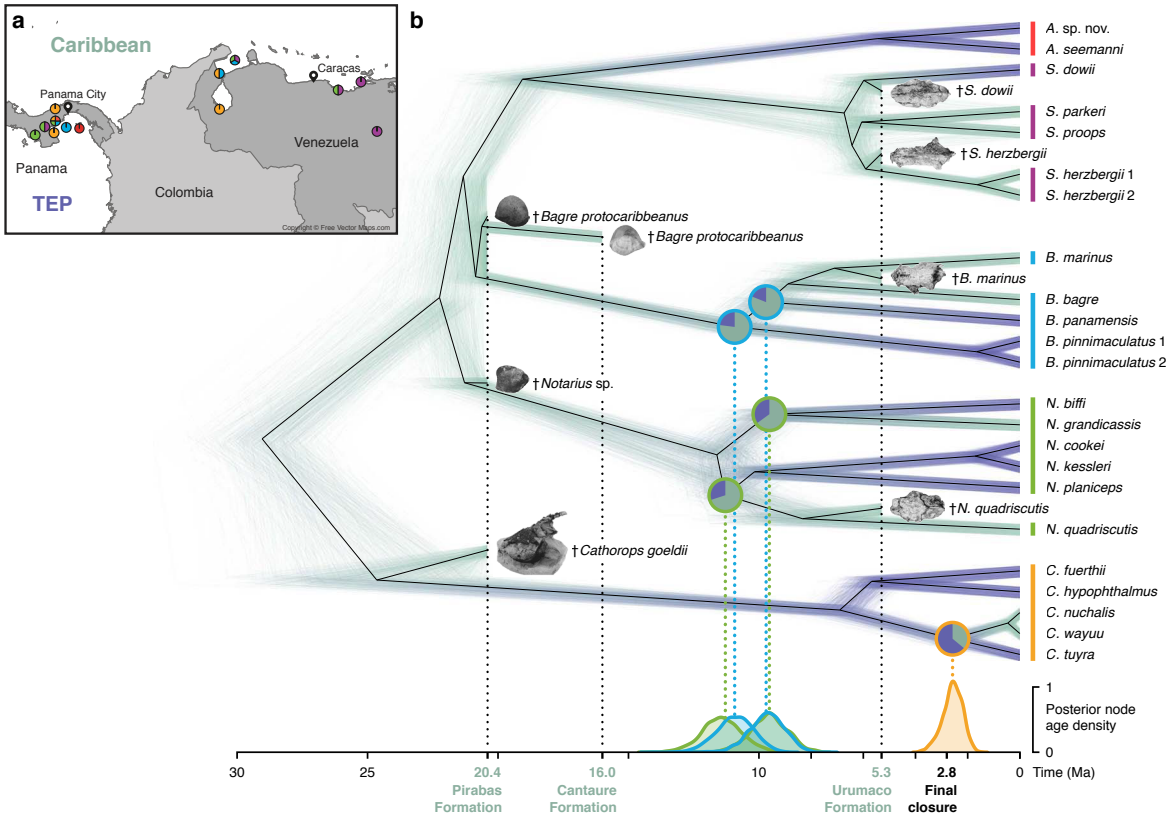


Figure 5: Time-calibrated species tree of Neotropical sea catfishes.
a) Map of Panama and north-western South America with sampling locations of specimens used in this study. Colors of circles indicate genera of specimens sampled at a location: *Ariopsis*, red; *Sciades*, purple; *Bagre*, blue; *Notarius*, green; *Cathorops*, orange. b) Posterior distribution of time-calibrated species trees inferred with SNAPP, with fossil taxa added a posteriori (images of otoliths and partial skulls are from Aguilera and de Aguilera 2004b and from Aguilera et al. 2013, 2014; see Table 7). Branch color indicates reconstructed geography: Caribbean; green, or Tropical Eastern Pacific (TEP); dark blue. Posterior densities of divergence times between Caribbean and Pacific lineages within *Notarius* (green), *Bagre* (blue), and *Cathorops* (orange) are shown below the species tree. Note that two divergence events around 10 Ma have nearly identical posterior density distributions: the divergence between *N. grandicassis* and *N. biffi* and the divergence between *B. panamensis* and the ancestor of *B. bagre* and *B. marinus*. Pie charts on nodes corresponding to divergences between Caribbean and Pacific lineages indicate posterior probabilities of ancestral distributions. All posterior estimates of node support, divergence times, and ancestral geography are summarized in Supplementary Table S4.

Discussion

Divergence-Time Estimation with Genome-Wide SNP Data

Our analyses based on simulated SNP data have demonstrated that SNAPP, combined with a molecular clock model, allows the precise and unbiased estimation of divergence times in the presence of incomplete lineage sorting. As expected, the precision of estimates increased with the number of SNPs used for the analysis. With 3 000 SNPs, the largest number of simulated SNPs used in our analyses, uncertainty in divergence times resulted almost exclusively from the width of the calibration density (Fig. 1). In addition to data set size, the placement of the node-age calibration also had a strong effect on the precision of divergence-time estimates, which was markedly improved when the root node was calibrated instead of a younger node. This suggests that future studies employing divergence-time estimation with SNAPP should make use of constraints on the root node if these are available from the fossil record, from biogeographic scenarios, or from previously published time-calibrated phylogenies (as in our analyses of empirical SNP data of Neotropical army ants and sea catfishes). While we did not test the performance of multiple calibration points with simulated data, the use of additional calibration points can be expected to further improve the precision of divergence-time estimates.

Perhaps more surprisingly, the computational time required for SNAPP analyses did not always increase with larger SNP numbers (Fig. 2). Instead, the time requirements even decreased with larger SNP numbers when younger nodes were calibrated instead of the root node, because less MCMC iterations were required for convergence. With calibrations on the root node, computational time increased with SNP numbers, but the increase was less-than-linear, most likely because the number of unique site patterns, which is highly correlated with the time required per iteration, showed signs of saturation in the larger

simulated SNP data sets. This suggests that further extensions to the SNP data may have relatively little impact on computational time requirements since many of the site patterns included in these extensions will already be present in the data set. Nevertheless, the addition of these SNPs will further improve the precision of divergence-time estimates, which implies that even analyses with extremely large numbers of SNPs may be worthwhile. However, how rapidly the set of unique site patterns approaches saturation likely varies among data sets and will depend on factors such as the number of species, the number of specimens per species, and the proportion of missing data.

An obvious limitation of our approach is the assumption of equal and constant population sizes on all branches of the phylogeny, which may be rarely met in practice. However, the linking of population sizes was necessary to achieve feasible run times for analyses of data sets with around 20 species (with this number of species, assuming an individual population size for each branch would require an additional 37 model parameters). The single population-size parameter estimated with our method will therefore most commonly represent an intermediate value within the range of the true population sizes of the taxa included in the data set. As a result, divergence times might be slightly overestimated for groups in which the population size is underestimated and vice versa. Nevertheless, we expect that the degree of this misestimation is minor compared to the bias introduced by the alternative strategy of concatenation (Fig. 4, Table 5), which is equivalent to the MSC model only when all population sizes are so small that incomplete lineage sorting is absent and all gene trees are identical in topology and branch lengths (Edwards et al. 2016).

Insights Into the Taxonomy of Neotropical Sea Catfishes

Different views on the taxonomy of sea catfishes (Ariidae) have been supported by phylogenetic inference based on morphological features (Marceniuk et al. 2012b) and

molecular data (Betancur-R. et al. 2007; Betancur-R. 2009). In the following, we address the most important differences between these views and how they are supported by our results, as well as new findings with regard to cryptic species.

Bagre and Cathorops.— The morphology-based phylogenetic analysis of Marceniuk et al. (2012b) supported an earlier proposal by Schultz (1944) to raise the genus *Bagre* to family status due to its extraordinary morphological distinctiveness and its inferred position outside of a clade combining almost all other genera of sea catfishes. On the other hand, molecular studies have recovered *Bagre* in a nested position within sea catfishes, a position that is also supported by our results (Betancur-R. et al. 2007, 2012; Betancur-R. 2009). The proposed status of *Bagre* as a separate family is therefore not supported by molecular data. Instead of *Bagre*, our phylogeny identified the genus *Cathorops* as the sister of a clade combining *Notarius*, *Bagre*, *Sciades*, and *Ariopsis*, in contrast not only to morphology-based analyses but also to previous molecular studies that recovered a clade combining *Cathorops*, *Bagre*, and *Notarius*, albeit with low support (Betancur-R. et al. 2007, 2012; Betancur-R. 2009).

Within the genus *Bagre*, cryptic speciation has previously been suggested in *B. pinnimaculatus* based on cranio-morphological differences and distinct mitochondrial haplotypes of populations from the Bay of Panama and from Rio Estero Salado, Panama (Stange et al. 2016). Our current results corroborate this hypothesis, given that the estimated divergence time of the two populations (*B. pinnimaculatus* 1 and *B. pinnimaculatus* 2 in Fig. 5) is old (1.66 Ma; 95% HPD: 2.30-1.08 Ma) compared to the expected coalescence time within a species
($T_{exp} = 2 \times Ng = 2 \times 127\,250 \times 2$ yr $= 509\,000$ yr; with N according to SNAPP's population size estimate and g according to an assumed generation time of two years for sea catfishes; Betancur-R. et al. 2008).

While *Cathorops nuchalis* has been declared a valid taxon based on morphological differentiation (Marceniuk et al. 2012a), mitochondrial sequences of this species were found to be indistinguishable from its sister species *C. wayuu* (Stange et al. 2016). In contrast, the nuclear SNP variation investigated here suggests that the two species are well differentiated and diverged 460 ka (95% HPD: 740-220 ka).

Notarius.— According to our results, *Notarius quadriscutis* is either the sister to a Pacific clade composed of *N. cookei*, *N. kessleri*, and *N. planiceps* (BPP: 0.54), the sister to *N. biffi* and *N. grandicassis* (BPP: 0.07), or the sister to all other sampled extant members of the genus (BPP: 0.39). Based on morphology, the species has previously been placed in genus *Aspistor* together with *N. luniscutis* and the extinct *N. verumquadriscutis* (Marceniuk et al. 2012b; Aguilera and Marceniuk 2012). However, molecular phylogenies have commonly recovered the species assigned to *Aspistor* nested within *Notarius* (Betancur-R. and Acero P. 2004; Betancur-R. et al. 2012) and thus do not support the distinction of the two genera. Regardless of the exact relationships of *Notarius quadriscutis* in our species tree, our analyses suggest that the lineage originated around the time of the crown divergence of *Notarius* (11.61 Ma; 95% HPD: 13.23-10.21 Ma) and is thus younger than the earliest fossils assigned to the genus, *Notarius* sp. (Early Miocene; Aguilera et al. 2014). This implies that considering *Aspistor* as separate from *Notarius* would also require a reevaluation of fossils assigned to *Notarius*.

Ariopsis and Sciades.— While molecular studies have supported the reciprocal monophyly of the genera *Ariopsis* and *Sciades* (Betancur-R. et al. 2007, 2012), species of genus *Ariopsis* appeared paraphyletic in the morphology-based phylogeny of Marceniuk et al. (2012b) and were there considered as members of *Sciades*. Our species tree inferred with SNAPP supports the results of previous molecular analyses since both genera appear as clearly monophyletic sister groups (BPP: 1.0) that diverged already in the Early Miocene

(19.06 Ma; 95% HPD: 20.94-17.45 Ma).

Within *Sciades*, differentiation of mitochondrial haplotypes has been observed between brackish-water and marine populations of *S. herzbergii* from Clarines, Venezuela, and from the Golf of Venezuela (Stange et al. 2016). Our relatively old divergence-time estimate (1.64 Ma; 95% HPD: 2.20-1.04 Ma) provides further support for substantial differentiation of the two populations (*S. herzbergii* 1 and *S. herzbergii* 2 in Fig. 5) that could be driven by ecological adaptations to their contrasting habitats.

Implications for the Closure of the Panamanian Isthmus

In agreement with our results based on simulated data, our reanalysis of genome-wide army ant data with both the MSC model and with concatenation indicated that recent divergence times can be overestimated if incomplete lineage sorting is not accounted for. As a result, the colonization of the North American landmass by army ants prior to the final closure of the Isthmus of Panama (2.8 Ma; O'Dea et al. 2016) was supported by our analyses using concatenation, but not by those using the MSC model. However, even the divergence times estimated with concatenation were generally younger than the divergence times reported by Winston et al. (2016), also on the basis of concatenation. This suggests that besides the variation introduced by the use of concatenation and the MSC, age estimates of army ant divergences were also influenced by other differences between our Bayesian divergence-time estimation and the analyses of Winston et al. (2016), which employed a penalized likelihood approach (Sanderson 2002) to estimate divergence times. These differences included not only the methodology used for time calibration, but also the number of specimens and alignment sites used in the analysis, as we had to filter the data set to comply with the assumption of the tree prior and to reduce the computational demands of the BEAST analysis. Nevertheless, our results

suggest that previous claims of army ant migration to the North American landmass prior to the final isthmus closure (Winston et al. 2016) should be viewed with caution.

By combining Bayesian phylogenetic inference with stochastic mapping of geographic distributions, our analyses of sea catfish SNP data allowed us to estimate the timing and the location of divergence events separating lineages of Caribbean and Pacific sea catfishes (Fig. 5). The youngest of these events is the divergence of the Caribbean common ancestor of *Cathorops nuchalis* and *C. wayuu* from the Pacific *C. tuyra*, which we estimated to have occurred around 2.58 Ma (95% HPD: 3.37-1.87 Ma). As this age estimate coincides with the final closure of the Panamanian Isthmus around 2.8 Ma (O'Dea et al. 2016; Groeneveld et al. 2014), it appears likely that the closure was causal for vicariant divergence within *Cathorops*. According to our reconstruction of ancestral geographic distributions, the common ancestor of the three species *C. nuchalis*, *C. wayuu*, and *C. tuyra* more likely lived in the TEP (BPP: 0.64) than in the Caribbean. We note that this discrete type of inference may appear incompatible with the assumption that these lineages speciated through vicariance, given that in this case, the geographic distribution of the common ancestor should have extended across both regions as long as they were still connected. While our discrete ancestral reconstructions did not allow us to model this scenario of partially continuous distributions explicitly, our reconstructions can be reconciled with it if the inferred discrete geography is viewed not as the exclusive distribution of a species, but as the center of its distribution instead.

Surprisingly, the divergence of Caribbean and Pacific lineages within *Cathorops* was the only splitting event in our sample of sea catfishes that could be associated with the final closure of the Panamanian Isthmus around 2.8 Ma, even though the closure could be expected to affect a large number of species simultaneously. Instead, near-simultaneous divergence events between Caribbean and Pacific lineages were inferred at a much earlier time, about 10 Ma, in the genera *Bagre* and *Notarius*. Within *Notarius*, *N. grandicassis* of

the Caribbean and the West Atlantic diverged from *N. biffi* of the TEP around 9.63 Ma (95% HPD: 10.99-8.30 Ma). This event may have coincided with the separation of Caribbean and Pacific lineages within *Bagre* (9.70 Ma; 95% HPD: 11.05-8.50 Ma), where the Pacific species *B. panamensis* diverged from a predominantly Caribbean (BPP: 0.81) ancestor that later gave rise to *B. bagre* and *B. marinus*. Two further divergence events between Caribbean and Pacific lineages of *Bagre* and *Notarius* were inferred slightly earlier, around 11 Ma. At 10.93 Ma (95% HPD: 12.29-9.60 Ma), the Pacific species *Bagre pinnimaculatus* diverged from the common ancestor of *B. marinus*, *B. bagre*, and *B. panamensis*, which likely had a distribution centered in the Caribbean (BPP: 0.77). Additionally, the common ancestor of the Pacific clade comprising *Notarius cookei*, *N. kessleri*, and *N. planiceps* diverged from the predominantly Caribbean (BPP: 0.70) lineage leading to *N. quadriscutis* at 11.29 Ma (95% HPD: 12.75-9.86 Ma).

Our time-calibrated species tree with reconstructed ancestral distributions (Fig. 5) shows further divergence events that separated Caribbean and Pacific lineages. The two sampled species of *Ariopsis* both occur in the TEP and diverged at about 19.06 Ma (95% HPD: 20.94-17.45 Ma) from the predominantly Caribbean genus *Sciades*. However, since *Ariopsis* also contains Caribbean species that we did not include in our data set, it remains unclear when and how often transitions between the Caribbean and the TEP took place in this genus. Caribbean origins of the genus *Cathorops* and of the species *Sciades dowii* are suggested by fossils from the Pirabas and Urumaco formations and indicate that these two lineages migrated to the Pacific after or simultaneous to the divergence from the fossil representatives. But since these divergence times were not estimated in our SNAPP analysis, the timing of migration of *Cathorops* and *Sciades dowii* also remains uncertain.

Regardless of these uncertainties, the near-simultaneous occurrence of several divergence events between Pacific and Caribbean lineages around 11-10 Ma suggests that geological processes associated with the emergence of the Panamanian Isthmus promoted

vicariance long before the final closure of the isthmus around 2.8 Ma. Thus, even though our reanalysis of Neotropical army ant data suggested that army ants did not colonize the North American landmass before the final isthmus closure, our results based on sea catfish data add to the body of molecular biological evidence that indicates the emergence of temporary land bridges in the Late Miocene, leading to the separation of marine populations and migration of terrestrial animals (Donaldson and Wilson Jr 1999; Musilová et al. 2008; Bacon et al. 2015a,b; Carrillo et al. 2015; Acero P. et al. 2016; Huang 2016) long before the Great American Biotic Interchange (Woodburne 2010). While Miocene land bridges have been supported by a number of studies (Collins et al. 1996; Montes et al. 2015; Bacon et al. 2015a), it remains debated whether all of the connections between the Caribbean and the Pacific closed prior to 2.8 Ma, and whether they were blocked at the same time (O'Dea et al. 2016). Nevertheless, even if land bridges did not block all passages simultaneously, their emergence might have disrupted the distributions of catfish populations if these were localized in areas away from the remaining openings.

Although the rapid succession of divergence events between Caribbean and Pacific sea catfish lineages around 11-10 Ma indicates vicariance as the result of emerging land bridges, we cannot exclude that these events were driven by other forms of speciation, such as ecological speciation, and that their clustering within this relatively short period is coincidential. To discriminate between these possible explanations, a better understanding of the ecology of the diverging taxa will be important. In addition, the compilation of further diversification timelines for other groups of marine Neotropical species may strengthen the support for vicariance if divergences in these groups were found to cluster around the same times as in sea catfishes. As our results based on simulations suggest, these future analyses may benefit from genome-wide SNP data, however, concatenation should be avoided in favor of the MSC model to produce the most accurate estimates of divergence times. Importantly, our results clearly demonstrate that regardless of the causes

of splitting events around 11-10 Ma, divergences between Caribbean and Pacific taxa are not necessarily linked to the final closure of the Panamanian Isthmus around 2.8 Ma. Thus, we reiterate earlier conclusions (Bacon et al. 2015a; De Baets et al. 2016) that the time of the final closure of the isthmus should no longer be used as a biogeographic calibration point for divergence-time estimation.

Conclusion

We have demonstrated that the software SNAPP, combined with a molecular clock model, allows highly precise and accurate divergence-time estimation based on SNP data and the multi-species coalescent model. Our method thus provides molecular biologists with a powerful tool to investigate the timing of recent divergence events with genome-wide data. Our application of this method to two genomic data sets of Neotropical army ants and sea catfishes led to mixed support for the suggested closure of the Isthmus of Panama in the Miocene. We showed that army ants of the genus *Eciton* may have colonized the North American landmass only after the final closure of the Isthmus around 2.8 Ma and that previous conclusions supporting Miocene and Pliocene colonization events may have been influenced by branch-length bias resulting from concatenation. In contrast, we identify a series of divergence events around 10 Ma between sea catfishes of the Caribbean and the TEP, which lends support to the hypothesis of Miocene isthmus closure. The rigorous application of divergence-time estimation with the multi-species coalescent model in future studies based on genomic data promises to contribute conclusive evidence for the timing and the effect of the emergence of the Panamanian Isthmus, one of the most significant events in recent geological history.

Funding

M.S. was funded by a Forschungskredit from the University of Zurich (FK-15-092); M.R.S.V. and W.S. were supported by the Swiss National Science Foundation Sinergia (Sinergia Grant CRSII3_136293). W.S. and M.M. were further supported by a grant from the European Research Council (CoG "CICHLID~X") awarded to W.S.

Supplementary Material

Supplementary Material, including figures, tables, and input and output files of SNAPP and BEAST can be found in the Dryad Data Repository XXX: XXX. Code for all analyses is provided on https://github.com/mmatschiner/panama, and the script "snapp_prep.rb" to generate SNAPP input files in XML format is available on https://github.com/mmatschiner/snapp_prep.

Acknowledgements

We thank Remco Bouckaert for advice on divergence-time estimation with SNAPP and Emiliano Trucchi for testing our method. We are grateful to Richard Cooke for hospitality and assistance in coordinating our fieldwork in Panama, Aureliano Valencia, and Máximo Jiménez for help with sampling in Panama and Tito Barros, Cathy Villalba, and Jorge Domingo Carrillo-Briceño for field assistance and paperwork in Venezuela. We acknowledge help from Max Winston and Corrie Moreau to reuse their data set, and we thank Alexander Leow for contribution to SNP analyses. Part of this work was performed using the Abel high-performance computing facilies (www.hpc.uio.no) of the University of Oslo and the sciCORE (http://scicore.unibas.ch) scientific computing core facility at the University of Basel.

*

References

Acero P., A., J. J. Tavera, R. Anguila, and L. Hernández. 2016. A new southern Caribbean species of angel shark (Chondrichthyes, Squaliformes, Squatinidae), including phylogeny and tempo of diversification of American species. Copeia 104:577–585.

Aguilera, O. and D. R. de Aguilera. 2004a. Amphi-American Neogene sea catfishes (Siluriformes, Ariidae) from northern South America. Spec. Pap. Paleontol. 71:29–48.

Aguilera, O. and D. R. de Aguilera. 2004b. New Miocene otolith-based scianid species (Pisces, Perciformes) from Venezuela. Spec. Pap. Paleontol. 71:49–59.

Aguilera, O. and A. P. Marceniuk. 2012. *Aspistor verumquadriscutis*, a new fossil species of sea catfishes (Siluriformes; Ariidae) from the upper Miocene of Venezuela. Swiss J. Palaeontol. 131:265–274.

Aguilera, O., W. Schwarzhans, H. Moraes-Santos, and A. Nepomuceno. 2014. Before the flood: Miocene otoliths from eastern Amazon Pirabas Formation reveal a Caribbean-type fish fauna. J. South Am. Earth Sci. 56:422–446.

Aguilera, O. A., H. Moraes-Santos, S. Costa, F. Ohe, C. Jaramillo, and A. Nogueira. 2013. Ariid sea catfishes from the coeval Pirabas (Northeastern Brazil), Cantaure, Castillo (Northwestern Venezuela), and Castilletes (North Colombia) formations (early Miocene), with description of three new species. Swiss J. Palaeontol. 132:45–68.

Aguilera, O. A., W. Schwarzhans, and P. Béarez. 2016. Otoliths of the Sciaenidae from the Neogene of tropical America. Palaeo Ichthyologica 14:1–128.

Alfaro, M. E., F. Santini, C. D. Brock, H. Alamillo, A. Dornburg, D. L. Rabosky, G. Carnevale, and L. J. Harmon. 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proc. Natl. Acad. Sci. USA 106:13410–13414.

Ali, J. R. and M. Huber. 2010. Mammalian biodiversity on Madagascar controlled by ocean currents. Nature 463:653–656.

Angelis, K. and M. dos Reis. 2015. The impact of ancestral population size and incomplete lineage sorting on Bayesian estimation of species divergence times. Curr. Zool. 61:874–885.

Bacon, C. D., P. Molnar, A. Antonelli, A. J. Crawford, C. Montes, and M. C. Vallejo-Pareja. 2016. Quaternary glaciation and the Great American Biotic Interchange. Geology 44:375–378.

Bacon, C. D., D. Silvestro, C. Jaramillo, B. T. Smith, P. Chakrabarty, and A. Antonelli. 2015a. Biological evidence supports an early and complex emergence of the Isthmus of Panama. Proc. Natl. Acad. Sci. USA 112:6110–6115.

Bacon, C. D., D. Silvestro, C. Jaramillo, B. T. Smith, P. Chakrabarty, and A. Antonelli. 2015b. Reply to Lessios and Marko et al.: Early and progressive migration across the Isthmus of Panama is robust to missing data and biases. Proc. Natl. Acad. Sci. USA 112:E5767–E5768.

Berghoff, S. M., D. J. C. Kronauer, K. J. Edwards, and N. R. Franks. 2008. Dispersal and population structure of a New World predator, the army ant *Eciton burchellii*. J. Evol. Biol. 21:1125–1132.

Bermingham, E., S. S. McCafferty, and A. P. Martin. 1997. Fish biogeography and

molecular clocks: perspectives from the Panamanian Isthmus. Pages 113–128 *in* Molecular Systematics of Fishes (T. D. Kocher and C. A. Stepien, eds.). Academic Press, San Diego, USA.

Betancur-R., R. 2009. Molecular phylogenetics and evolutionary history of ariid catfishes revisited: a comprehensive sampling. BMC Evol. Biol. 9:175.

Betancur-R., R. and A. Acero P. 2004. Description of *Notarius biffi* n. sp. and redescription of *N. insculptus* (Jordan and Gilbert) (Siluriformes: Ariidae) from the Eastern Pacific, with evidence of monophyly and limits of *Notarius*. Zootaxa 703:1–20.

Betancur-R., R., A. Acero P., E. Bermingham, and R. Cooke. 2007. Systematics and biogeography of New World sea catfishes (Siluriformes: Ariidae) as inferred from mitochondrial, nuclear, and morphological evidence. Mol. Phylogenet. Evol. 45:339–357.

Betancur-R., R., A. P. Marceniuk, and P. Béarez. 2008. Taxonomic status and redescription of the Gillbacker sea catfish (Siluriformes: Ariidae: *Sciades parkeri*). Copeia 2008:827–834.

Betancur-R., R., G. Ortí, A. M. Stein, A. P. Marceniuk, and R. A. Pyron. 2012. Apparent signal of competition limiting diversification after ecological transitions from marine to freshwater habitats. Ecol. Lett. 15:822–830.

Bloch, J. I., E. D. Woodruff, A. R. Wood, A. F. Rincon, A. R. Harrington, G. S. Morgan, D. A. Foster, C. Montes, C. A. Jaramillo, N. A. Jud, D. S. Jones, and B. J. MacFadden. 2016. First North American fossil monkey and early Miocene tropical biotic interchange. Nature 533:243–246.

Bouckaert, R. R. 2010. DensiTree: making sense of sets of phylogenetic trees. Bioinformatics 26:1372–1373.

Bouckaert, R. R., J. Heled, D. Kühnert, T. Vaughan, C.-H. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. J. Drummond. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLOS Comput. Biol. 10:e1003537.

Brady, S. G., B. L. Fisher, T. R. Schultz, and P. S. Ward. 2014. The rise of army ants and their relatives: diversification of specialized predatory doryline ants. BMC Evol. Biol. 14:93.

Bryant, D., R. R. Bouckaert, J. Felsenstein, N. A. Rosenberg, and A. RoyChoudhury. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Mol. Biol. Evol. 29:1917–1932.

Campbell, K. E., D. R. Prothero, L. Romero-Pittman, F. Hertel, and N. Rivera. 2010. Amazonian magnetostratigraphy: dating the first pulse of the Great American Faunal Interchange. J. South Am. Earth Sci. 29:619–626.

Carranza-Castañeda, O. and W. E. Miller. 2004. Late Tertiary terrestrial mammals from central Mexico and their relationship to South American immigrants. Rev. Bras. Paleontolog. 7:249–261.

Carrillo, J. D., A. Forasiepi, C. Jaramillo, and M. R. Sánchez-Villagra. 2015. Neotropical mammal diversity and the Great American Biotic Interchange: spatial and temporal variation in South America's fossil record. Front. Genet. 5:451.

Carrillo-Briceño, J. D., O. A. Aguilera, C. De Gracia, G. Aguirre-Fernández, R. Kindlimann, and M. R. Sánchez-Villagra. 2016. An Early Neogene Elasmobranch fauna from the southern Caribbean (Western Venezuela). Palaeontol. Electron. 19.2.27A:1–32.

Catchen, J. M., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait. 2011. Stacks: building and genotyping loci de novo from short-read sequences. G3 1:171–182.

Cervigón, F., R. Cipriani, W. Fischer, L. Garibaldi, M. Hendrickx, A. J. Lemus, M. R, J. M. Poutiers, G. Robaina, and B. Rodriquez. 1993. Field Guide to the Commercial Marine and Brackish-water Resources of the Northern Coast of South America. FAO, Rome, Italy.

Chifman, J. and L. Kubatko. 2014. Quartet inference from SNP data under the coalescent model. Bioinformatics 30:3317–3324.

Coates, A. G., L. S. Collins, M. P. Aubry, and W. A. Berggren. 2004. The Geology of the Darien, Panama, and the late Miocene-Pliocene collision of the Panama arc with northwestern South America. Geol. Soc. Am. Bull. 116:1327–1344.

Coates, A. G. and R. F. Stallard. 2013. How old is the Isthmus of Panama? Bull. Mar. Sci. 89:801–813.

Cohen, K., S. Finney, P. Gibbard, and J.-X. Fan. 2013. The ICS International Chronostratigraphic Chart. Episodes 36: 199-204. Available from `http://www.stratigraphy.org/ICSchart/ChronostratChart2015-01.pdf`.

Collins, L. S., A. G. Coates, W. A. Berggren, M. P. Aubry, and J. Zhang. 1996. The late Miocene Panama isthmian strait. Geology 24:687–690.

Cox, J. C., J. E. J. Ingersoll, and S. A. Ross. 1985. A theory of the term structure of interest rates. Econometrica 53:1–25.

De Baets, K., A. Antonelli, and P. C. J. Donoghue. 2016. Tectonic blocks and molecular clocks. Phil. Trans. R. Soc. B 371:20160098.

de Queiroz, A. 2014. The Monkey's Voyage: How Improbable Journeys Shaped the History of Life. Basic Books, Perseus Books Group, New York, USA.

Demos, T. C., J. C. Kerbis Peterhans, T. A. Joseph, J. D. Robinson, B. Agwanda, and M. J. Hickerson. 2015. Comparative population genomics of African montane forest mammals support population persistence across a climatic gradient and Quaternary climatic cycles. PLOS ONE 10:e0131800.

Donaldson, K. A. and R. R. Wilson Jr. 1999. Amphi-Panamic geminates of snook (Percoidei: Centropomidae) provide a calibration of the divergence rate in the mitochondrial DNA control region of fishes. Mol. Phylogenet. Evol. 13:208–213.

Drummond, A. J. and R. R. Bouckaert. 2015. Bayesian evolutionary analysis with BEAST 2. Cambridge University Press, Cambridge, UK.

Drummond, A. J., G. K. Nicholls, A. G. Rodrigo, and W. Solomon. 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161:1307–1320.

Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29:1969–1973.

Eaton, D. A. R. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics 30:1844–1849.

Edwards, S. V., Z. Xi, A. Janke, B. C. Faircloth, J. E. McCormack, T. C. Glenn, B. Zhong, S. Wu, E. M. Lemmon, A. R. Lemmon, A. D. Leaché, L. Liu, and C. C. Davis. 2016. Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. Mol. Phylogenet. Evol. 94:447–462.

Farris, D. W., C. Jaramillo, G. Bayona, S. A. Restrepo-Moreno, C. Montes, A. Cardona, A. Mora, R. J. Speakman, M. D. Glascock, and V. Valencia. 2011. Fracturing of the Panamanian Isthmus during initial collision with South America. Geology 39:1007–1010.

Felsenstein, J. 1992. Estimating effective population size from samples of sequences: inefficiency of pairwise and segregating sites as compared to phylogenetic estimates. Genet. Res. 59:139–147.

Flynn, J. J., B. J. Kowallis, C. Nuñez, O. Carranza-Castañeda, W. E. Miller, C. C. Swisher, III, and E. Lindsay. 2005. Geochronology of Hemphillian-Blancan aged strata, Guanajuato, Mexico, and implications for timing of the Great American Biotic Interchange. J. Geol. 113:287–307.

Frailey, C. D. and K. E. Campbell. 2012. Two new genera of peccaries (Mammalia, Artiodactyla, Tayassuidae) from upper Miocene deposits of the Amazon Basin. J. Paleont. 86:852–877.

Gatesy, J. and M. S. Springer. 2013. Concatenation versus coalescence versus "concatalescence". Proc. Natl. Acad. Sci. USA 110:E1179–E1179.

Gatesy, J. and M. S. Springer. 2014. Phylogenetic analysis at deep timescales: Unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. Mol. Phylogenet. Evol. 80:231–266.

Gavryushkina, A., D. Welch, T. Stadler, and A. J. Drummond. 2014. Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. PLOS Comput. Biol. 10:e1003919.

Gregory, T. 2016. Animal Genome Size Database. `http://www.genomesize.com`.

Groeneveld, J., E. C. Hathorne, S. Steinke, H. DeBey, A. Mackensen, and R. Tiedemann. 2014. Glacial induced closure of the Panamanian Gateway during Marine Isotope Stages (MIS) 95–100 (~2.5 Ma). Earth Planet Sci. Lett. 404:296–306.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22:160–174.

Heath, T. A., J. P. Huelsenbeck, and T. Stadler. 2014. The fossilized birth-death process for coherent calibration of divergence-time estimates. Proc. Natl. Acad. Sci. USA 111:E2957–E2966.

Heled, J. and R. R. Bouckaert. 2013. Looking for trees in the forest: summary tree from posterior samples. BMC Evol. Biol. 13:1–11.

Heled, J. and A. J. Drummond. 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27:570–580.

Hirschfeld, S. E. 1968. Plio-Pleistocene megalonychid sloths of North America. Bull. Florida State Mus. 12:213–296.

Hobolth, A., O. F. Christensen, T. Mailund, and M. H. Schierup. 2007. Genomic relationships and speciation times of human, chimpanzee, and gorilla inferred from a coalescent hidden Markov model. PLOS Genet. 3:e7.

Huang, J.-P. 2016. The great American biotic interchange and diversification history in *Dynastes* beetles (Scarabaeidae; Dynastinae). Zool. J. Linnean Soc. 178:88–96.

Huelsenbeck, J. P., R. Nielsen, and J. P. Bollback. 2003. Stochastic mapping of morphological characters. Syst. Biol. 52:131–158.

Jackson, J. and A. O'Dea. 2013. Timing of the oceanographic and biological isolation of the Caribbean Sea from the tropical eastern Pacific Ocean. Bull. Mar. Sci. 89:779–800.

Jukes, T. H. and C. R. Cantor. 1969. Evolution of protein molecules. Pages 21–132 *in* Mammalian Protein Metabolism (H. N. Munro, ed.). Academic Press, New York.

Kay, R. F. 2015. Biogeography in deep time – What do phylogenetics, geology, and paleoclimate tell us about early platyrrhine evolution? Mol. Phylogenet. Evol. 82:358–374.

Kubatko, L. S., B. C. Carstens, and L. L. Knowles. 2009. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. Bioinformatics 25:971–973.

Kubatko, L. S. and J. H. Degnan. 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56:17–24.

Lanier, H. C. and L. L. Knowles. 2012. Is recombination a problem for species-tree analyses? Syst. Biol. 61:691–701.

Laurito, C. A. and A. L. Valerio. 2012. Primer registro fósil de *Pliometanastes* sp. (Mammalia, Xenarthra, Megalonychidae) para el Mioceno Superior de Costa Rica, América Central. Una nueva pista en la comprensión del Pre-GABI. Rev. Geol. Amér. Central 47:95–108.

Lepage, T., D. Bryant, H. Philippe, and N. Lartillot. 2007. A general comparison of relaxed molecular clock models. Mol. Biol. Evol. 24:2669–2680.

Lessios, H. A. 2008. The great American schism: divergence of marine organisms after the rise of the Central American Isthmus. Annu. Rev. Ecol. Evol. Syst. 39:63–91.

Linkem, C. W., V. N. Minin, and A. D. Leaché. 2016. Detecting the anomaly zone in species trees and evidence for a misleading signal in higher-level skink phylogeny (Squamata: Scincidae). Syst. Biol. 65:465–477.

Lischer, H. E. L., L. Excoffier, and G. Heckel. 2014. Ignoring heterozygous sites biases phylogenomic estimates of divergence times: implications for the evolutionary history of *Microtus* voles. Mol. Biol. Evol. 31:817–831.

Liu, L. 2008. BEST: Bayesian estimation of species trees under the coalescent model. Bioinformatics 24:2542–2543.

Liu, L., L. Yu, and S. V. Edwards. 2010. A maximum pseudo-likelihood approach for estimating species trees under the coalescent. BMC Evol. Biol. 10:302.

Lozier, J. D., J. M. Jackson, M. E. Dillon, and J. P. Strange. 2016. Population genomics of divergence among extreme and intermediate color forms in a polymorphic insect. Ecol. Evol. 6:1075–1091.

Maddison, W. P. 1997. Gene trees in species trees. Syst. Biol. 46:523–536.

Marceniuk, A. P., R. Betancur-R., A. Acero P., and J. Muriel-Cunha. 2012a. Review of the genus *Cathorops* (Siluriformes: Ariidae) from the Caribbean and Atlantic South America, with description of a new species. Copeia 2012:77–97.

Marceniuk, A. P., N. A. Menezes, and M. R. Britto. 2012b. Phylogenetic analysis of the family Ariidae (Ostariophysi: Siluriformes), with a hypothesis on the monophyly and relationships of the genera. Zool. J. Linnean Soc. 165:534–669.

Marshall, L. G. 1988. Land mammals and the Great American Interchange. Am. Sci. 76:380–388.

Matschiner, M., Z. Musilová, J. M. I. Barth, Z. Starostová, W. Salzburger, M. Steel, and R. R. Bouckaert. 2016. Bayesian phylogenetic estimation of clade ages supports trans-Atlantic dispersal of cichlid fishes. Syst. Biol. Advance Access, doi: 10.1093/sysbio/syv076.

McCormack, J. E., J. Heled, K. S. Delaney, A. T. Peterson, and L. L. Knowles. 2011. Calibrating divergence times on species trees versus gene trees: implications for speciation history of *Aphelocoma* jays. Evolution 65:184–202.

Mendes, F. K. and M. W. Hahn. 2016. Gene tree discordance causes apparent substitution rate variation. Syst. Biol. 65:711–721.

Meunier, F. J. 2012. Skeletochronological studies of cyclical growth of freshwater fishes in French Guiana. A review. Cybium 36:55–62.

Meyer, B. S., M. Matschiner, and W. Salzburger. 2016. Disentangling incomplete lineage sorting and introgression to refine species-tree estimates for Lake Tanganyika cichlid fishes. Syst. Biol. Advance Access, doi: 10.1093/sysbio/syv069.

Mirarab, S. and T. Warnow. 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics 31:i44–52.

Montes, C., G. Bayona, A. Cardona, D. M. Buchs, C. A. Silva, S. Morón, N. Hoyos, D. A. Ramirez, C. A. Jaramillo, and V. Valencia. 2012. Arc-continent collision and orocline formation: Closing of the Central American seaway. J. Geophys. Res. 117:B04105.

Montes, C., A. Cardona, C. Jaramillo, A. Pardo, J. C. Silva, V. Valencia, C. Ayala, L. C. Pérez-Angel, L. A. Rodriguez-Parra, V. Ramirez, and H. Niño. 2015. Middle Miocene closure of the Central American Seaway. Science 348:226–229.

Musilová, Z., O. Říčan, K. Janko, and J. Novák. 2008. Molecular phylogeny and biogeography of the Neotropical cichlid fish tribe Cichlasomatini (Teleostei: Cichlidae: Cichlasomatinae). Mol. Phylogenet. Evol. 46:659–672.

O'Dea, A., H. A. Lessios, A. G. Coates, R. I. Eytan, S. A. Restrepo-Moreno, A. L. Cione, L. S. Collins, A. de Queiroz, D. W. Farris, R. D. Norris, R. F. Stallard, M. O.

Woodburne, O. Aguilera, M.-P. Aubry, W. A. Berggren, A. F. Budd, M. A. Cozzuol, S. E. Coppard, H. Duque-Caro, S. Finnegan, G. M. Gasparini, E. L. Grossman, K. G. Johnson, L. D. Keigwin, N. Knowlton, E. G. Leigh, J. S. Leonard-Pingel, P. B. Marko, N. D. Pyenson, P. G. Rachello-Dolmen, E. Soibelzon, L. Soibelzon, J. A. Todd, G. J. Vermeij, and J. B. C. Jackson. 2016. Formation of the Isthmus of Panama. Sci. Adv. 2:e1600883.

Ogilvie, H. A., R. R. Bouckaert, and A. J. Drummond. 2016a. StarBEAST2 brings faster species tree inference and accurate estimates of substitution rates. bioRxiv preprint, doi: 10.1101/038455.

Ogilvie, H. A., J. Heled, D. Xie, and A. J. Drummond. 2016b. Computational performance and statistical accuracy of *BEAST and comparisons with other methods. Syst. Biol. Advance Access, doi: 10.1093/sysbio/syv118.

Osborne, A. H., D. R. Newkirk, J. Groeneveld, E. E. Martin, R. Tiedemann, and M. Frank. 2014. The seawater neodymium and lead isotope record of the final stages of Central American Seaway closure. Paleoceanography 29:715–729.

Plummer, M., N. Best, K. Cowles, and K. Vines. 2006. CODA: Convergence diagnosis and output analysis for MCMC. R News 6:7–11.

Prothero, D. R., K. E. Campbell, Jr, B. L. Beatty, and C. D. Frailey. 2014. New late Miocene dromomerycine artiodactyl from the Amazon Basin: implications for interchange dynamics. J. Paleont. 88:434–443.

Rabosky, D. L., F. Santini, J. Eastman, S. A. Smith, B. Sidlauskas, J. Chang, and M. E. Alfaro. 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. Nat. Commun. 4:1958.

Rambaut, A. and N. C. Grassly. 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Comput. Appl. Biosci. 13:235–238.

Rambaut, A., M. Suchard, D. Xie, and A. Drummond. 2014. Tracer v1.6. Available from `http://beast.bio.ed.ac.uk/Tracer`.

Rannala, B. and Z. Yang. 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics 164:1645–1656.

Revell, L. J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). Method Ecol. Evol. 3:217–223.

Roch, S. and M. Steel. 2014. Likelihood-based tree reconstruction on a concatenation of aligned sequence data sets can be statistically inconsistent. Theor. Popul. Biol. 100:56–62.

Roesti, M., A. P. Hendry, W. Salzburger, and D. Berner. 2012. Genome divergence during evolutionary diversification as revealed in replicate lake-stream stickleback population pairs. Mol. Ecol. 21:2852–2862.

RoyChoudhury, A. and E. A. Thompson. 2012. Ascertainment correction for a population tree via a pruning algorithm for likelihood computation. Theor. Popul. Biol. 82:59–65.

Ru, D., K. Mao, L. Zhang, X. Wang, Z. Lu, and Y. Sun. 2016. Genomic evidence for polyphyletic origins and interlineage gene flow within complex taxa: a case study of *Picea brachytyla* in the Qinghai-Tibet Plateau. Mol. Ecol. 25:2373–2386.

Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19:101–109.

Savin, S. M. and R. G. Douglas. 1985. Sea level, climate, and the Central American land bridge. Pages 303–324 *in* The Great American Biotic Interchange (F. G. Stehli and S. D. Webb, eds.). Plenum Press, New York, USA.

Scally, A., J. Y. Dutheil, L. W. Hillier, G. E. Jordan, I. Goodhead, J. Herrero, A. Hobolth, T. Lappalainen, T. Mailund, T. Marquès-Bonet, S. McCarthy, S. H. Montgomery, P. C. Schwalie, Y. A. Tang, M. C. Ward, Y. Xue, B. Yngvadottir, C. Alkan, L. N. Andersen, Q. Ayub, E. V. Ball, K. Beal, B. J. Bradley, Y. Chen, C. M. Clee, S. Fitzgerald, T. A. Graves, Y. Gu, P. Heath, A. Heger, E. Karakoc, A. Kolb-Kokocinski, G. K. Laird, G. Lunter, S. Meader, M. Mort, J. C. Mullikin, K. Munch, T. D. O'Connor, A. D. Phillips, J. Prado-Martinez, A. S. Rogers, S. Sajjadian, D. Schmidt, K. Shaw, J. T. Simpson, P. D. Stenson, D. J. Turner, L. Vigilant, A. J. Vilella, W. Whitener, B. Zhu, D. N. Cooper, P. de Jong, E. T. Dermitzakis, E. E. Eichler, P. Flicek, N. Goldman, N. I. Mundy, Z. Ning, D. T. Odom, C. P. Ponting, M. A. Quail, O. A. Ryder, S. M. Searle, W. C. Warren, R. K. Wilson, M. H. Schierup, J. Rogers, C. Tyler-Smith, and R. Durbin. 2012. Insights into hominid evolution from the gorilla genome sequence. Nature 483:169–175.

Schultz, L. P. 1944. The catfishes of Venezuela, with descriptions of thirty-eight new forms. Proceedings of the United States National Museum 94:173–338.

Scornavacca, C. and N. Galtier. 2016. Incomplete lineage sorting in mammalian phylogenomics. Syst. Biol. Advance Access, doi:10.1093/sysbio/syw082.

Sepulchre, P., T. Arsouze, Y. Donnadieu, C. Jaramillo, J. Le Bras, E. Martin, C. Montes, and A. J. Waite. 2014. Consequences of shoaling of the Central American Seaway determined from modeling Nd isotopes. Paleoceanography 29:176–189.

Springer, M. S. and J. Gatesy. 2016. The gene tree delusion. Mol. Phylogenet. Evol. 94:1–33.

Stange, M., G. Aguirre-Fernández, R. G. Cooke, T. Barros, W. Salzburger, and M. R. Sánchez-Villagra. 2016. Evolution of opercle bone shape along a macrohabitat gradient: species identification using mtDNA and geometric morphometric analyses in neotropical sea catfishes (Ariidae). Ecol. Evol. 6:5817–5830.

Suh, A., L. Smeds, and H. Ellegren. 2015. The dynamics of incomplete lineage sorting across the ancient adaptive radiation of Neoavian birds. PLOS Biol. 13:e1002224.

Sukumaran, J. and M. T. Holder. 2010. DendroPy: a Python library for phylogenetic computing. Bioinformatics 26:1569–1571.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures on Mathematics in the Life Sciences 17:57–86.

Winston, M. E., D. J. C. Kronauer, and C. S. Moreau. 2016. Early and dynamic colonization of Central America drives speciation in Neotropical army ants. Mol. Ecol. Advance Access, doi:10.1111/mec.13846.

Woodburne, M. O. 2010. The Great American Biotic Interchange: dispersals, tectonics, climate, sea level and holding pens. J. Mamm. Evol. 17:245–264.

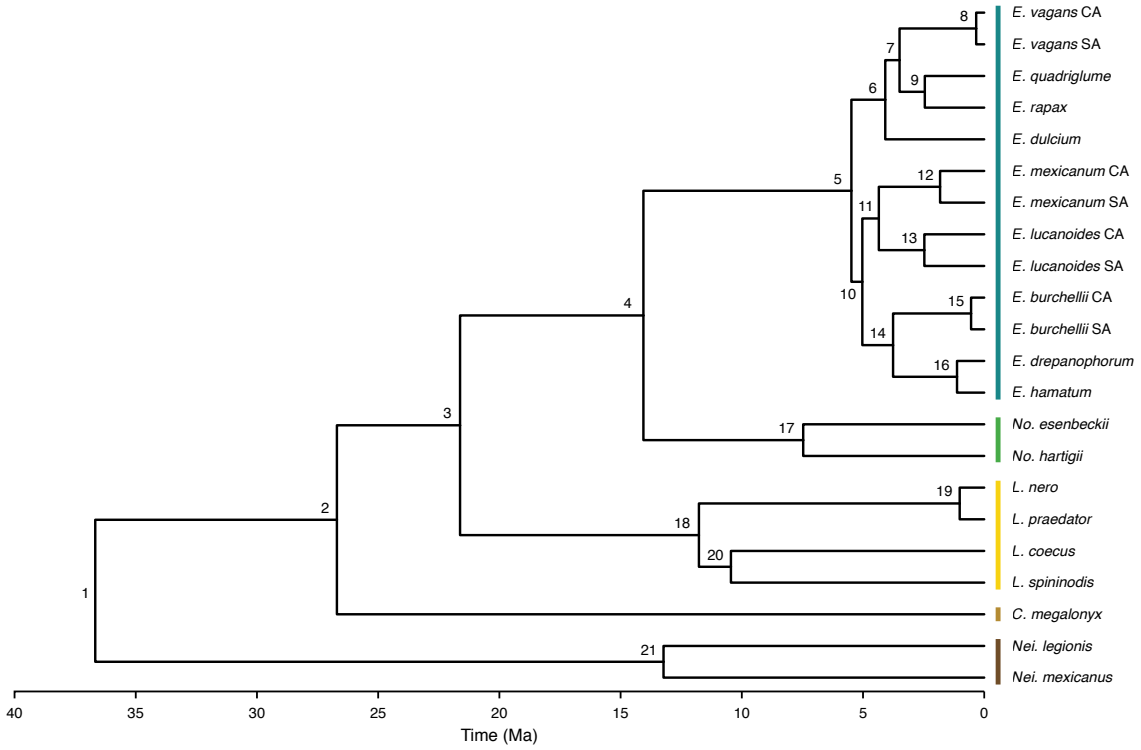
Yule, G. U. 1925. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. Phil. Trans. R. Soc. B 213:21–87.

Chapter 2

Bayesian divergence-time estimation with genome-wide SNP data of sea catfishes (Ariidae) Supports Miocene closure of the Panamanian Isthmus

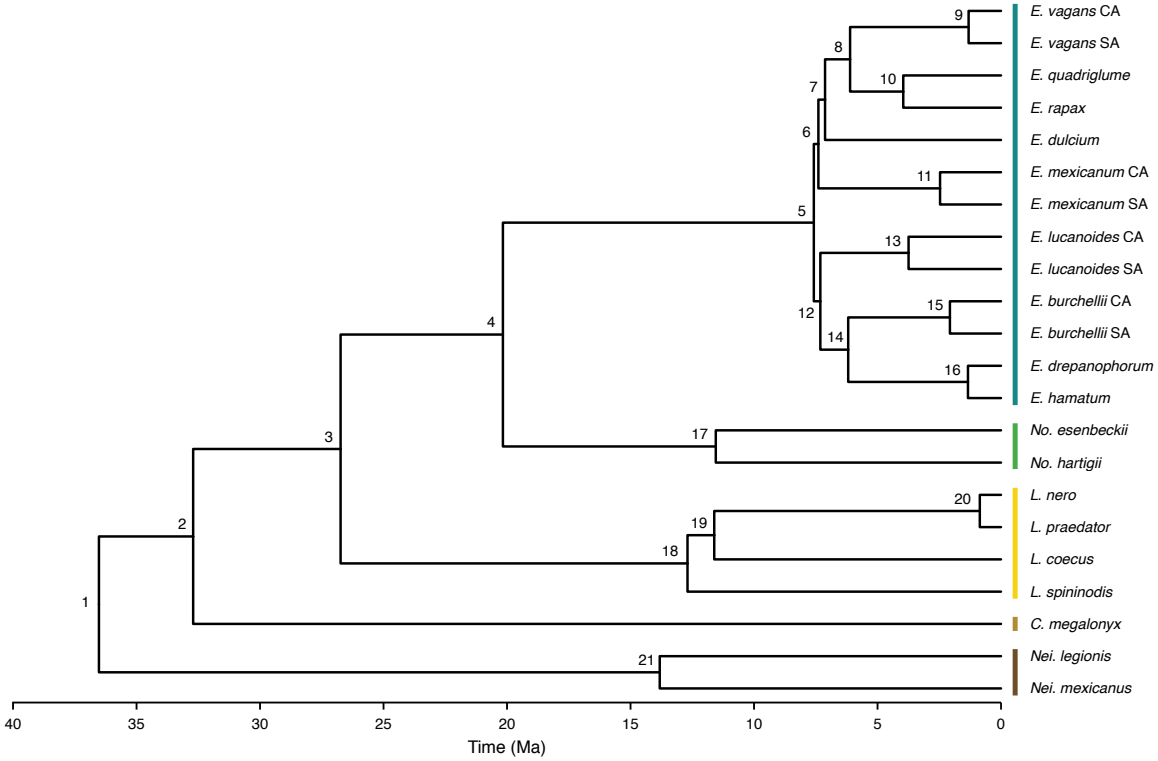
Supplememtary material

Supplementary Figure S1: Maximum clade credibilty tree of Neotropical army ants based on the MSC model.
Node ids are provided to allow comparison to node support and age estimates given in Supplementary Table S2. Colors indicates genera: *Eciton*, blue-green; *Nomamyrmex*, green; *Labidus*, yellow; *Cheliomyrmex*, light brown; *Neivamyrmex*, dark brown. CA, Central America; SA, South America.



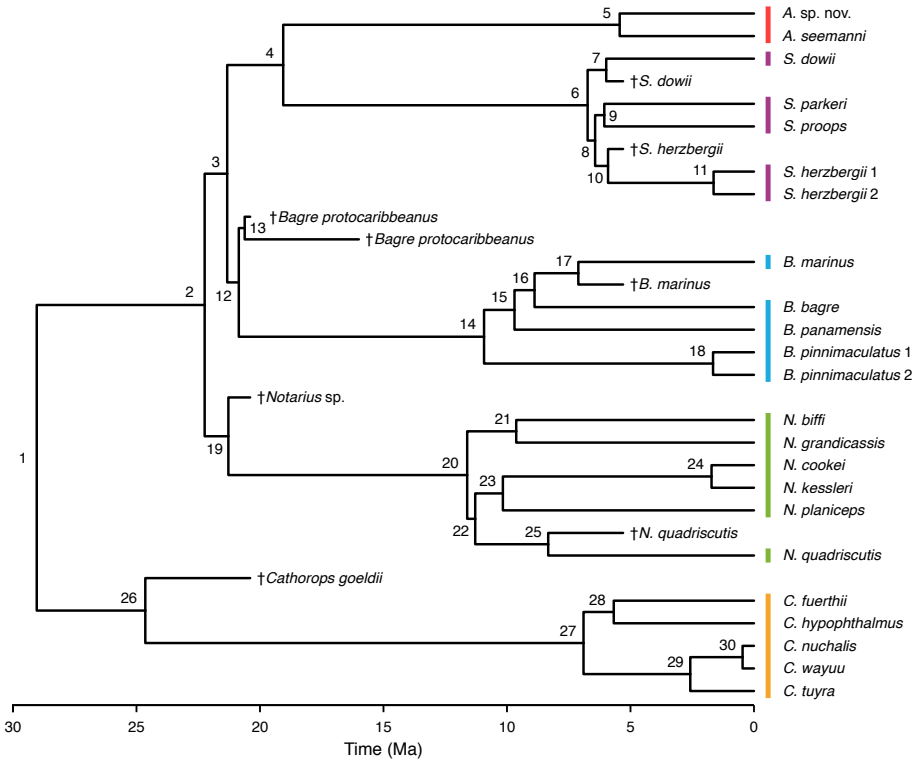
Supplementary Figure S2: Maximum clade credibilty tree of Neotropical army ants based on concatenation.

Node ids are provided to allow comparison to node support and age estimates given in Supplementary Table S3. Colors indicates genera as in Supplementary Figure S1: *Eciton*, blue-green; *Nomamyrmex*, green; *Labidus*, yellow; *Cheliomyrmex*, light brown; *Neivamyrmex*, dark brown. CA, Central America; SA, South America.



Supplementary Figure S3: Maximum clade credibilty tree of Neotropical sea catfishes based on the MSC model.

Node ids are provided to allow comparison to node support, age estimates, and reconstructed ancestral geography given in Supplementary Table S4. Colors indicates genera as in Figure 5 and in Stange et al. (2016): *Ariopsis*, red; *Sciades*, purple; *Bagre*, blue; *Notarius*, green; *Cathorops*, orange.



Supplementary Table S1: Studies using SNAPP, published before December 2016.

Study	Analysis type	# groups	# individuals	#loci
Bell et al. (2015)	phylogeny	4	21	467 SNPs
Blanca et al. (2015)	phylogeny	16	16	2 313 SNPs
Boucher et al. (2016)	BFD*	1-4	90	1 449-1 665 SNPs
Bryant et al. (2012)	phylogeny	6	69	1 997 AFLP
Bryson et al. (2016)	phylogeny	9	54	312 SNPs
Burns and Tsurusaki (2016)	phylogeny	5	20	226 SNPs
Card et al. (2016)	BFD*	2-3	33	1 686 SNPs
Chifman and Kubatko (2014)	phylogeny	10	20	1 027 026 SNPs
Clucas et al. (2016)	phylogeny	4	8	2 626-2 668 SNPs
DaCosta and Sorenson (2016)	phylogeny	10-15	10-15	2 856-3 098 SNPs
Demos et al. (2015)	phylogeny	6-7	28-31	52 580-56 194 SNPs
Foote and Morin (2016)	phylogeny	5	43	1 346 SNPs
Ford et al. (2015)	phylogeny	11	44	1 266 SNPs
Fraser et al. (2016)	BFD*	3	73	75 712 SNPs
Giarla and Esselstyn (2015)	phylogeny	9-19	19	1 170 SNPs
Gohli et al. (2014)	phylogeny	10	59	166-2 000 SNPs
Gottscho et al. (2017)	BFD*	2-6	64	468 SNPs
Gratton et al. (2016)	BFD*	4-6	46	2 039 SNPs
Hamlin and Arnold (2014)	phylogeny	6-12	43-67	560-1 140 SNPs
Harvey and Brumfield (2015)	phylogeny	5-8	72	3 379 SNPs
Herrera and Shank (2016)	phylogeny, BFD*	3-24	31	1 203 SNPs
Kautt et al. (2016)	phylogeny	7	28	1 290-1 772 SNPs
Lambert et al. (2013)	phylogeny	4	138	821 AFLPs
Leaché et al. (2014)	BFD*	2-5	46	129-1 087 SNPs
Li et al. (2015)	phylogeny, BFD*	4-10	10	1 041 SNPs
Lischer et al. (2014)	phylogeny	4	15	1 552 SNPs
Lozier et al. (2016)	phylogeny	5	41	1 568 SNPs
MacLeod et al. (2015)	phylogeny	12	33	6 257 SNPs
De Maio et al. (2015)	phylogeny	4-8	4-80	3-1 000 genes
Manthey et al. (2015)	phylogeny	8	41	13 421 SNPs
Manthey et al. (2016)	phylogeny	16	30	605-1 128 SNPs
Mason and Taylor (2015)	BFD*	2-4	20	35-1 587 SNPs
McCormack et al. (2015)	phylogeny	26	26	1 388 SNPs
Meik et al. (2015)	phylogeny	5	26	2 409 SNPs
Morin et al. (2015)	phylogeny	12	113	42 SNPs
Mrinalini et al. (2015)	BFD*	2-4	50	298 AFLPs
Nater et al. (2015)	phylogeny	4	48	16 000 SNPs
Nicotra et al. (2016)	phylogeny	8	23	463 SNPs
Nieto-Montes de Oca et al. (2017)	BFD*	1-4	50	401-430 SNPs
Ogilvie et al. (2016)	phylogeny	8-12	11-45	791-5 907 SNPs
Olšavská et al. (2016)	phylogeny	14	120	442 AFLPs
Oswald et al. (2016)	BFD*	2-3	13	14 285 SNPs
Paun et al. (2016)	phylogeny	26	79	1 506 SNPs
Potter et al. (2016)	phylogeny, BFD*	3-5	8-14	2 084 SNPs
Razkin et al. (2016)	phylogeny, BFD*	7-10	25	368-875 SNPs

BFD*, Bayes Factor delimitation with genomic data (Leaché et al. 2014)

Supplementary Table S1 (continued):

Study	Analysis type	# groups	# individuals	#loci
Rheindt et al. (2014)	phylogeny	4-5	12	947-954 SNPs
Rittmeyer and Austin (2015)	BFD*	2-3	12	941-1 973 SNPs
Sovic et al. (2016)	phylogeny, BFD*	2-5	12-15	1 099 SNPs
Stervander et al. (2016)	phylogeny	6-7	12	1 590-26 629 SNPs
Stervander et al. (2015)	phylogeny	16	16	3 421 SNPs
Streicher et al. (2014)	phylogeny	9	39	353 SNPs
Takahashi et al. (2016)	phylogeny	10	60	390 AFLPs
Takahashi et al. (2015)	phylogeny	8	90	522 AFLPs
Tremetsberger et al. (2016)	phylogeny	4	36	207 AFLPs
Winger et al. (2015)	phylogeny	6	14	1 767 SNPs
Yoder et al. (2016)	phylogeny	6	29	3 986-4 613 SNPs
Yuan et al. (2016)	phylogeny	6	6	2 921 SNPs
Zarza et al. (2016)	phylogeny	24	24	2 501 SNPs

BFD*, Bayes Factor delimitation with genomic data (Leaché et al. 2014)

Supplementary Table S2: Node support and age estimates for the species tree of Neotropical army ants based on the MSC model. Lower and upper boundaries given for age estimates describe 95% HPD intervals. Node ids correspond to those given in Supplementary Figure S1.

		Age estimates (Ma)		
Node	Node support (BPP)	Mean	Upper	Lower
1	1.00	36.67	45.76	27.66
2	1.00	26.70	34.98	19.25
3	1.00	21.62	28.27	15.27
4	1.00	14.06	18.51	9.30
5	1.00	5.48	7.52	3.52
6	0.98	4.08	5.86	2.44
7	0.79	3.49	5.10	2.05
8	1.00	0.33	0.71	0.05
9	0.99	2.45	3.73	1.28
10	0.51	5.03	6.81	3.22
11	0.78	4.35	6.13	2.71
12	1.00	1.82	3.02	0.76
13	1.00	2.47	3.88	1.22
14	1.00	3.76	5.40	2.19
15	1.00	0.54	1.12	0.13
16	1.00	1.12	1.95	0.37
17	1.00	7.47	10.61	4.54
18	1.00	11.77	15.81	7.53
19	1.00	1.01	2.02	0.16
20	0.64	10.45	14.33	6.65
21	1.00	13.23	18.02	8.71

Supplementary Table S3: Node support and age estimates for the species tree of Neotropical army ants based on concatenation. Lower and upper boundaries given for age estimates describe 95% HPD intervals. Node ids correspond to those given in Supplementary Figure S2.

		Age estimates (Ma)		
Node	Node support (BPP)	Mean	Upper	Lower
1	1.00	36.52	45.23	27.38
2	1.00	32.71	40.67	24.70
3	1.00	26.74	33.24	20.19
4	1.00	20.17	25.10	15.25
5	1.00	7.58	9.44	5.74
6	1.00	7.39	9.19	5.58
7	1.00	7.12	8.84	5.37
8	1.00	6.11	7.59	4.61
9	1.00	1.31	1.65	1.00
10	1.00	3.96	4.91	2.98
11	1.00	2.47	3.09	1.88
12	1.00	7.32	9.13	5.55
13	1.00	3.74	4.68	2.83
14	1.00	6.19	7.74	4.71
15	1.00	2.07	2.56	1.55
16	1.00	1.34	1.66	1.01
17	1.00	11.55	14.33	8.69
18	1.00	12.69	15.83	9.63
19	1.00	11.61	14.42	8.73
20	1.00	0.86	1.07	0.64
21	1.00	13.82	17.19	10.42

Supplementary Table S4: Node support, age estimates, and reconstructed ancestral geography for the species tree of Neotropical sea catfishes based on the MSC model. Lower and upper boundaries given for age estimates describe 95% HPD intervals. Node ids correspond to those given in Supplementary Figure S3.

		Age estimates (Ma)			Geography (BPP)	
Node	Node support (BPP)	Mean	Upper	Lower	Caribbean	TEP
1	1.00	29.04	31.35	26.86	0.81	0.19
2	1.00	22.24	24.06	20.68	0.99	0.01
3	0.92	21.33	23.00	20.40	1.00	0.00
4	1.00	19.06	20.94	17.45	0.89	0.11
5	1.00	5.44	6.63	4.39	0.11	0.89
6	1.00	6.74	7.90	5.74	1.00	0.00
7	1.00	5.98	6.98	5.30	1.00	0.00
8	0.62	6.43	7.50	5.48	1.00	0.00
9	0.79	6.06	7.08	4.98	1.00	0.00
10	1.00	5.91	6.73	5.30	1.00	0.00
11	1.00	1.64	2.20	1.04	1.00	0.00
12	1.00	20.86	21.86	20.40	1.00	0.00
13	1.00	20.62	21.28	20.40	1.00	0.00
14	1.00	10.93	12.29	9.60	0.77	0.23
15	1.00	9.70	11.05	8.50	0.81	0.19
16	0.98	8.89	10.17	7.67	0.89	0.11
17	1.00	7.11	9.01	5.31	0.97	0.03
18	1.00	1.66	2.30	1.08	0.02	0.98
19	1.00	21.28	22.78	20.40	0.99	0.01
20	1.00	11.61	13.23	10.21	0.70	0.30
21	1.00	9.63	10.99	8.30	0.65	0.35
22	0.54	11.29	12.75	9.86	0.70	0.30
23	1.00	10.17	11.59	8.90	0.55	0.45
24	1.00	1.72	2.36	1.15	0.02	0.98
25	1.00	8.33	11.23	5.32	0.90	0.10
26	1.00	24.65	28.95	20.41	0.87	0.13
27	1.00	6.90	8.12	5.85	0.20	0.80
28	1.00	5.68	6.74	4.63	0.12	0.88
29	1.00	2.58	3.37	1.87	0.36	0.64
30	1.00	0.46	0.74	0.22	1.00	0.00

References

Bell, R. C., R. C. Drewes, and K. R. Zamudio. 2015. Reed frog diversification in the Gulf of Guinea: Overseas dispersal, the progression rule, and in situ speciation. Evolution 69:904–915.

Blanca, J., J. Montero-Pau, C. Sauvage, G. Bauchet, E. Illa, M. J. Díez, D. Francis, M. Causse, E. van der Knaap, and J. Cañizares. 2015. Genomic variation in tomato, from wild ancestors to contemporary breeding accessions. BMC Genomics 16:257.

Boucher, F. C., G. Casazza, P. Szövényi, and E. Conti. 2016. Sequence capture using RAD probes clarifies phylogenetic relationships and species boundaries in *Primula* sect. *Auricula*. Mol. Phylogenet. Evol. 104:60–72.

Bryant, D., R. R. Bouckaert, J. Felsenstein, N. A. Rosenberg, and A. RoyChoudhury. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Mol. Biol. Evol. 29:1917–1932.

Bryson, R. W., Jr., W. E. Savary, A. J. Zellmer, R. B. Bury, and J. E. McCormack. 2016. Genomic data reveal ancient microendemism in forest scorpions across the California Floristic Province. Mol. Ecol. 25:3731–3751.

Burns, M. and N. Tsurusaki. 2016. Male reproductive morphology across latitudinal clines and under long-term female sex-ratio bias. Integr. Comp. Biol. 56:715–727.

Card, D. C., D. R. Schield, R. H. Adams, A. B. Corbin, B. W. Perry, A. L. Andrew, G. I. M. Pasquesi, E. N. Smith, T. Jezkova, S. M. Boback, W. Booth, and T. A. Castoe. 2016. Phylogeographic and population genetic analyses reveal multiple species of *Boa* and independent origins of insular dwarfism. Mol. Phylogenet. Evol. 102:104–116.

Chifman, J. and L. Kubatko. 2014. Quartet inference from SNP data under the coalescent model. Bioinformatics 30:3317–3324.

Clucas, G. V., J. L. Younger, D. Kao, A. D. Rogers, J. Handley, G. D. Miller, P. Jouventin, P. Nolan, K. Gharbi, K. J. Miller, and T. Hart. 2016. Dispersal in the sub-Antarctic: king penguins show remarkably little population genetic differentiation across their range. BMC Evol. Biol. 16:211.

DaCosta, J. M. and M. D. Sorenson. 2016. ddRAD-seq phylogenetics based on nucleotide, indel, and presence–absence polymorphisms: Analyses of two avian genera with contrasting histories. Mol. Phylogenet. Evol. 94:122–135.

De Maio, N., D. Schrempf, and C. Kosiol. 2015. PoMo: An allele frequency-based approach for species tree estimation. Syst. Biol. 64:1018–1031.

Demos, T. C., J. C. Kerbis Peterhans, T. A. Joseph, J. D. Robinson, B. Agwanda, and M. J. Hickerson. 2015. Comparative population genomics of African montane forest mammals support population persistence across a climatic gradient and Quaternary climatic cycles. PLOS ONE 10:e0131800.

Foote, A. D. and P. A. Morin. 2016. Genome-wide SNP data suggest complex ancestry of sympatric North Pacific killer whale ecotypes. Heredity 117:316–325.

Ford, A. G. P., K. K. Dasmahapatra, L. Rüber, K. Gharbi, T. Cezard, and J. J. Day. 2015. High levels of interspecific gene flow in an endemic cichlid fish adaptive radiation from an extreme lake environment. Mol. Ecol. 24:3421–3440.

Fraser, C. I., A. Mcgaughran, A. Chuah, and J. M. Waters. 2016. The importance of replicating genomic analyses to verify phylogenetic signal for recently evolved lineages. Mol. Ecol. 25:3683–3695.

Giarla, T. C. and J. A. Esselstyn. 2015. The challenges of resolving a rapid, recent radiation: empirical and simulated phylogenomics of Philippine shrews. Syst. Biol. 64:727–740.

Gohli, J., E. H. Leder, E. Garcia-del Rey, L. E. Johannessen, A. Johnsen, T. Laskemoen, M. Popp, and J. T. Lifjeld. 2014. The evolutionary history of Afrocanarian blue tits inferred from genomewide SNPs. Mol. Ecol. 24:180–191.

Gottscho, A. D., D. A. Wood, A. G. Vandergast, J. Lemos-Espinal, J. Gatesy, and T. W. Reeder. 2017. Lineage diversification of fringe-toed lizards (Phrynosomatidae: *Uma notata* complex) in the Colorado Desert: Delimiting species in the presence of gene flow. Mol. Phylogenet. Evol. 106:103–117.

Gratton, P., E. Trucchi, A. Trasatti, G. Riccarducci, S. Marta, G. Allegrucci, D. Cesaroni, and V. Sbordoni. 2016. Testing classical species properties with contemporary data: how "bad species" in the brassy ringlets (*Erebia tyndarus* complex, Lepidoptera) turned good. Syst. Biol. 65:292–303.

Hamlin, J. A. P. and M. L. Arnold. 2014. Determining population structure and hybridization for two iris species. Ecol. Evol. 4:743–755.

Harvey, M. G. and R. T. Brumfield. 2015. Genomic variation in a widespread Neotropical bird (*Xenops minutus*) reveals divergence, population expansion, and gene flow. Mol. Phylogenet. Evol. 83:305–316.

Herrera, S. and T. M. Shank. 2016. RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. Mol. Phylogenet. Evol. 100:70–79.

Kautt, A. F., G. Machado-Schiaffino, and A. Meyer. 2016. Multispecies outcomes of sympatric speciation after admixture with the source population in two radiations of Nicaraguan crater lake cichlids. PLOS Genet. 12:e1006157.

Lambert, S. M., A. J. Geneva, D. Luke Mahler, and R. E. Glor. 2013. Using genomic data to revisit an early example of reproductive character displacement in Haitian *Anolis* lizards. Mol. Ecol. 22:3981–3995.

Leaché, A. D., M. K. Fujita, V. N. Minin, and R. R. Bouckaert. 2014. Species delimitation using genome-wide SNP data. Syst. Biol. 63:534–542.

Li, C., S. Corrigan, L. Yang, N. Straube, M. Harris, M. Hofreiter, W. T. White, and G. J. P. Naylor. 2015. DNA capture reveals transoceanic gene flow in endangered river sharks. Proc. Natl. Acad. Sci. USA 112:13302–13307.

Lischer, H. E. L., L. Excoffier, and G. Heckel. 2014. Ignoring heterozygous sites biases phylogenomic estimates of divergence times: implications for the evolutionary history of *Microtus* voles. Mol. Biol. Evol. 31:817–831.

Lozier, J. D., J. M. Jackson, M. E. Dillon, and J. P. Strange. 2016. Population genomics of divergence among extreme and intermediate color forms in a polymorphic insect. Ecol. Evol. 6:1075–1091.

MacLeod, A., A. Rodríguez, M. Vences, P. Orozco-terWengel, C. García, F. Trillmich, G. Gentile, A. Caccone, G. Quezada, and S. Steinfartz. 2015. Hybridization masks speciation in the evolutionary history of the Galápagos marine iguana. Proc. R. Soc. B 282:20150425.

Manthey, J. D., L. C. Campillo, K. J. Burns, and R. G. Moyle. 2016. Comparison of target-capture and restriction-site associated DNA sequencing for phylogenomics: a test in cardinalid tanagers (Aves, genus: *Piranga*). Syst. Biol. 65:640–650.

Manthey, J. D., J. Klicka, and G. M. Spellman. 2015. Chromosomal patterns of diversity and differentiation in creepers: a next-gen phylogeographic investigation of *Certhia americana*. Heredity 115:165–172.

Mason, N. A. and S. A. Taylor. 2015. Differentially expressed genes match bill morphology and plumage despite largely undifferentiated genomes in a Holarctic songbird. Mol. Ecol. 24:3009–3025.

McCormack, J. E., W. L. E. Tsai, and B. C. Faircloth. 2015. Sequence capture of ultraconserved elements from bird museum specimens. Mol. Ecol. Resour. 16:1189–1203.

Meik, J. M., J. W. Streicher, A. M. Lawing, O. Flores-Villela, and M. K. Fujita. 2015. Limitations of climatic data for inferring species boundaries: insights from speckled rattlesnakes. PLOS ONE 10:e0131435.

Morin, P. A., K. M. Parsons, F. I. Archer, M. C. Ávila-Arcos, L. G. Barrett-Lennard, L. Dalla Rosa, S. Duchêne, J. W. Durban, G. M. Ellis, S. H. Ferguson, J. K. Ford, M. J. Ford, C. Garilao, M. T. P. Gilbert, K. Kaschner, C. O. Matkin, S. D. Petersen, K. M. Robertson, I. N. Visser, P. R. Wade, S. Y. W. Ho, and A. D. Foote. 2015. Geographic and temporal dynamics of a global radiation and diversification in the killer whale. Mol. Ecol. 24:3964–3979.

Mrinalini, R. S. Thorpe, S. Creer, D. Lallias, L. Dawnay, B. L. Stuart, and A. Malhotra. 2015. Convergence of multiple markers and analysis methods defines the genetic distinctiveness of cryptic pitvipers. Mol. Phylogenet. Evol. 92:266–279.

Nater, A., R. Burri, T. Kawakami, L. Smeds, and H. Ellegren. 2015. Resolving evolutionary relationships in closely related species with whole-genome sequencing data. Syst. Biol. 64:1000–1017.

Nicotra, A. B., C. Chong, J. G. Bragg, C. R. Ong, N. C. Aitken, A. Chuah, B. Lepschi, and J. O. Borevitz. 2016. Population and phylogenomic decomposition via genotyping-by-sequencing in Australian *Pelargonium*. Mol. Ecol. 25:2000–2014.

Nieto-Montes de Oca, A., A. J. Barley, R. N. Meza-Lázaro, U. O. García-Vázquez, J. G. Zamora-Abrego, R. C. Thomson, and A. D. Leaché. 2017. Phylogenomics and species delimitation in the knob-scaled lizards of the genus *Xenosaurus* (Squamata: Xenosauridae) using ddRADseq data reveal a substantial underestimation of diversity. Mol. Phylogenet. Evol. 106:241–253.

Ogilvie, H. A., J. Heled, D. Xie, and A. J. Drummond. 2016. Computational performance and statistical accuracy of *BEAST and comparisons with other methods. Syst. Biol. Advance Access, doi: 10.1093/sysbio/syv118.

Olšavská, K., M. Slovák, K. Marhold, and E. Štubňová. 2016. On the origins of Balkan endemics: the complex evolutionary history of the *Cyanus napulifer* group (Asteraceae). Ann. Bot. 118:1071–1088.

Oswald, J. A., M. G. Harvey, R. C. Remsen, D. U. Foxworth, S. W. Cardiff, D. L. Dittmann, L. C. Megna, M. D. Carling, and R. T. Brumfield. 2016. Willet be one species or two? A genomic view of the evolutionary history of *Tringa semipalmata*. The Auk 133:593–614.

Paun, O., B. Turner, E. Trucchi, J. Munzinger, M. W. Chase, and R. Samuel. 2016. Processes driving the adaptive radiation of a tropical tree (*Diospyros*, Ebenaceae) in New Caledonia, a biodiversity hotspot. Syst. Biol. 65:212–227.

Potter, S., J. G. Bragg, B. M. Peter, K. Bi, and C. Moritz. 2016. Phylogenomics at the tips: inferring lineages and their demographic history in a tropical lizard, *Carlia amax*. Mol. Ecol. 25:1367–1380.

Razkin, O., G. Sonet, K. Breugelmans, M. J. Madeira, B. J. Gómez-Moliner, and T. Backeljau. 2016. Species limits, interspecific hybridization and phylogeny in the cryptic land snail complex *Pyramidula*: The power of RADseq data. Mol. Phylogenet. Evol. 101:267–278.

Rheindt, F. E., M. K. Fujita, P. R. Wilton, and S. V. Edwards. 2014. Introgression and phenotypic assimilation in *Zimmerius* flycatchers (Tyrannidae): population genetic and phylogenetic inferences from genome-wide SNPs. Syst. Biol. 63:134–152.

Rittmeyer, E. N. and C. C. Austin. 2015. Combined next-generation sequencing and morphology reveal fine-scale speciation in Crocodile Skinks (Squamata: Scincidae: *Tribolonotus*). Mol. Ecol. 24:466–483.

Sovic, M. G., A. C. Fries, and H. L. Gibbs. 2016. Origin of a cryptic lineage in a threatened reptile through isolation and historical hybridization. Heredity 117:358–366.

Stange, M., G. Aguirre-Fernández, R. G. Cooke, T. Barros, W. Salzburger, and M. R. Sánchez-Villagra. 2016. Evolution of opercle bone shape along a macrohabitat gradient: species identification using mtDNA and geometric morphometric analyses in neotropical sea catfishes (Ariidae). Ecol. Evol. 6:5817–5830.

Stervander, M., P. Alström, U. Olsson, U. Ottosson, B. Hansson, and S. Bensch. 2016. Multiple instances of paraphyletic species and cryptic taxa revealed by mitochondrial and nuclear RAD data for *Calandrella* larks (Aves: Alaudidae). Mol. Phylogenet. Evol. 102:233–245.

Stervander, M., J. C. Illera, L. Kvist, P. Barbosa, N. P. Keehnen, P. Pruisscher, S. Bensch, and B. Hansson. 2015. Disentangling the complex evolutionary history of the Western Palearctic blue tits (*Cyanistes* spp.) - phylogenomic analyses suggest radiation by multiple colonization events and subsequent isolation. Mol. Ecol. 24:2477–2494.

Streicher, J. W., T. J. Devitt, C. S. Goldberg, J. H. Malone, H. Blackmon, and M. K. Fujita. 2014. Diversification and asymmetrical gene flow across time and space: lineage sorting and hybridization in polytypic barking frogs. Mol. Ecol. 23:3273–3291.

Takahashi, H., T. Kondou, N. Takeshita, T.-H. Hsu, and M. Nishida. 2016. Evolutionary process of iwame, a markless form of the red-spotted masu salmon *Oncorhynchus masou ishikawae*, in the Ono River, Kyushu. Ichthyol. Res. 63:132–144.

Takahashi, H., N. Takeshita, H. Tanoue, S. Ueda, H. Takeshima, T. Komatsu, I. Kinoshita, and M. Nishida. 2015. Severely depleted genetic diversity and population structure of a large predatory marine fish (*Lates japonicus*) endemic to Japan. Conserv. Genet. 16:1155–1165.

Tremetsberger, K., M. Á. Ortiz, A. Terrab, F. Balao, R. Casimiro-Soriguer, M. Talavera, and S. Talavera. 2016. Phylogeography above the species level for perennial species in a composite genus. AoB PLANTS 8:plv142.

Winger, B. M., P. A. Hosner, G. A. Bravo, A. M. Cuervo, N. Aristizábal, L. E. Cueto, and J. M. Bates. 2015. Inferring speciation history in the Andes with reduced-representation sequence data: an example in the bay-backed antpittas (Aves; Grallariidae; *Grallaria hypoleuca* s. l.). Mol. Ecol. 24:6256–6277.

Yoder, A. D., C. R. Campbell, M. B. Blanco, M. dos Reis, J. U. Ganzhorn, S. M. Goodman, K. E. Hunnicutt, P. A. Larsen, P. M. Kappeler, R. M. Rasoloarison, J. M. Ralison, D. L. Swofford, and D. W. Weisrock. 2016. Geogenetic patterns in mouse lemurs (genus *Microcebus*) reveal the ghosts of Madagascar's forests past. Proc. Natl. Acad. Sci. USA 113:8049–8056.

Yuan, H., J. Jiang, F. A. Jiménez, E. P. Hoberg, J. A. Cook, K. E. Galbreath, and C. Li. 2016. Target gene enrichment in the cyclophyllidean cestodes, the most diverse group of tapeworms. Mol. Ecol. Resour. 16:1095–1106.

Zarza, E., B. C. Faircloth, W. L. E. Tsai, R. W. Bryson, Jr., J. Klicka, and J. E. McCormack. 2016. Hidden histories of gene flow in highland birds revealed with genomic markers. Mol. Ecol. 25:5144–5157.

Chapter 3

Phylomorphospace and morphological diversity of northern Neotropical Ariidae reveals conservatism among genera despite habitat transitions

Phylomorphospace and morphological diversity of northern Neotropical Ariidae reveals conservatism among genera despite macorhabitat transitions

Madlen Stange[1,*], Gabriel Aguirre-Fernández[1], Walter Salzburger[2], Marcelo R. Sánchez-Villagra[1]

[1] Palaeontological Institute and Museum, University of Zurich, Karl-Schmid-Strasse 4, 8006 Zurich, Switzerland

[2] Zoological Institute, University of Basel, Vesalgasse 1, 4051 Basel, Switzerland

[*] corresponding author

Abstract

Background

Morphological convergence triggered by trophic adaptations is a common pattern in adaptive radiations. The study of shape in an evolutionary context is usually restricted to well-studied fish models. We take advantage of the recently revised systematics of New World Ariidae and investigate skull shape evolution in six genera of northern Neotropical Ariidae. They constitute a lineage that diversified in the marine habitat but repeatedly adapted to freshwater habitats. 3D geometric morphometrics was applied for the first time in catfish skulls and phylogenetically informed statistical analyses were performed to test for the impact of habitat on skull diversification after habitat transition in this lineage.

Results

We found that skull shape is conserved throughout phylogeny. A morphospace analysis revealed that freshwater and marine species exhibit similar Procrustes variances, and brackish species revealed constrained group-specific shape variances. Yet freshwater species occupy the smallest shape space compared to marine and brackish species (based on partial disparity), and marine and freshwater species have the largest distance between each other. The genus *Bagre* occupies a distinct shape space in phylomorphospace, whereas *Ariopsis* mostly overlaps with other genera.

Conclusions

The morphospace analysis highlights once more the morphological distinctiveness of *Bagre*, and explains the conflicting classification of *Ariopsis* and *Sciades* species on the basis of morphological inference. However, molecular data clearly separates the two genera and also shows that *Bagre* is phylogenetically nested within the other investigated genera. The morphospace analysis shows that conservatism dominates skull shape evolution among genera although some phylogenetic lability can be observed within genera.

Keywords: disparity, fish, geometric morphometrics, morphological evolution, otolith, phylogeny

Background

Convergent evolution has been shown to be common among adaptive radiations, as in sticklebacks [1–3], African Lake cichlids [4], Midas cichlids [5], or African barbs [6]. Morphological convergence that is triggered by ecological convergence is typically manifested in a feeding-associated feature, the skull. Its shape evolution has been studied in a variety of teleost fishes to determine the factors that influence evolutionary change. This has been done from two main perspectives. One is a developmental perspective, by examining factors as modularity and integration. Morphological evolution is constrained by development and integration and can be enhanced by modularity [7] but this is not a universal pattern [8]. Several recent studies on teleost fishes investigated whether integration or modularity facilitate radiation [9–12]. The other approach examines species diversification from an adaptational perspective by investigating factors as predator avoidance, niche occupation, or ecological function. These studies either focused on the biomechanical link of skull or mandible shape to functional ecology [13–17] or explicitly investigated convergent evolution of skull shape and biotic and abiotic covariates [6, 16, 18, 19]. Other studies are exploratory or descriptive in nature [20–22]. Some of the above mentioned studies revealed that species that inhabit the same ecological niche converge in shape [9, 13–15, 17, 18]. This seems to be a common pattern also among terrestrial vertebrates [23–26], although mismatches of form and function triggered by behavioural plasticity and diverse constraints exist as well [27]. In terrestrial vertebrates the impact of phylogenetic dependence on shape similarity among closely related taxa has been tested explicitly [26, 28, 29] and several studies have examined these in teleost fishes, as well, but only using individual adaptive traits for ecological niches [30–33]. So far, statistical tests for the significance of phylogenetic dependence on the composite adaptive trait of skull shape has only been conducted in a single study on triggerfishes (Balistidae), revealing that phylogeny is the major explanatory factor of similarity in skull shape [16].

Previous studies that have investigated skull shape diversity and evolution in teleost fishes either applied traditional morphometric approaches based on linear measurements [13, 14, 34], or - more commonly - landmark-based two-dimensional geometric morphometrics (2D GM) [6, 10, 11, 15–17, 20–22, 35]. Currently, there is only a single study on fish skull shape evolution that applies more sophisticated three-dimensional geometric morphometric analyses (3D GM) [9].

In this study we aimed to add to the spectrum of methods that are applied to study teleost skull shape diversity and to fill an "organismal gap". Firstly, we focus on co-variation of skull shape and habitat in a teleost group understudied with respect to their natural history, Ariidae. By doing so we aimed at identifying causes of shape variation that accompany the

morphological diversification of ariid species from a defined geographic area during a macrohabitat transition. Secondly, we combine the study of shape variation of a composite adaptive trait, the teleost skull, using 3D GM, representing the second study only to use this method in teleosts. We consider the analysis of 3D shape being advantageous as it, unlike 2D GM, avoids artefacts that result from positioning of not perfectly flat objects. Thirdly, we combine the study of shape variation with a test for the influence of phylogenetic dependence and take that dependence into account while analysing co-variation with habitat. Although many studies are concerned with the impact of ecological or trophic niche and convergence of skull shape or function, the phylogenetic context of the species under study are rarely taken into account.

We analysed skull shape variation in 28 species representing six genera (*Ariopsis*, *Bagre*, *Cathorops*, *Notarius*, *Potamarius*, and *Sciades*) of the non-monophyletic group of northern Neotropical Ariidae (sea catfishes) from marine, brackish, and freshwater habitats. Ariidae are widely distributed in all tropical and subtropical marine regions, as well as in near-coastal rivers and lakes. Freshwater environments are inhabited by species that adapted secondarily to freshwater [36] during independent habitat transitions [37]. Only 4 % of ray-finned fishes manage to live in both marine and freshwater [38]. The evolutionary history of habitat transitions from freshwater (ariid ancestors) to marine and back to freshwater, with the availability of intermediate species with brackish occurrence, makes the Ariidae a valid system to study marine-freshwater transitions and associated shape changes. We categorise the ecological co-variate in macrohabitats (marine, brackish, and freshwater) as a collective proxy for differences in their ecological niche. Very little is known about the natural history of the individual species, e.g. feeding preferences, migration behaviour during breeding season, number of growth cycles per year, age at maturity, or longevity. We focused on the geographically circumvented taxa of the northern Neotropics in order to concentrate our resources and to appropriately identify the habitat that the species and different populations of the same species occupy. We make use of a new time-calibrated phylogenetic hypothesis of northern Neotropical Ariidae that was inferred by Bayesian inference from single-nucleotide polymorphism (SNP) markers, which also includes two newly discovered cryptic species [39]. Furthermore, by exploring taxa in and around the Panamanian Isthmus, we can address the potential influence of the separation of the Pacific and Atlantic Ocean in Neogene times. This is a topical matter for which new data and analyses are needed to test alternative hypotheses on the tempo and mode of the formation of the Isthmus and of evolution around it [39].

We hypothesized that a) skull shape in northern Neotropical Ariidae is similar within the same macro-habitats, driven by e.g., differences in osmoregulation or biotic resources affecting freshwater, brackish, and marine living species, differently; b) skull shape should consequently differ among macrohabitats; c)

rates of shape evolution is higher in Tropical Eastern Pacific (TEP) species compared to Caribbean species due to differences in ocean productivity [40] and also higher in marine species compared to freshwater species due to the younger age of the derived freshwater species. Furthermore, we compare morphospace occupancy on genus-level to examine similarities caused by phylogenetic dependencies.

We investigated skull shape diversification in northern Neotropical Ariidae from marine, brackish, and freshwater habitats using 3D geometric morphometrics and corrected the shape data for phylogenetic dependence of species. We found that skull shape is strongly affected by phylogenetic dependence. Procrustes analyses show that although fresh and marine species exhibit similar Procrustes variances, freshwater species occupy the smallest shape space compared to brackish and marine species. This result is robust also after the exclusion of Bagre (pure marine) and Potamarius (pure freshwater) to account for possible confounding effects of these habitat-specialized genera. Ariopsis fully overlaps with other genera in morphospace, complicating morphology-based methods for species delimitation.

Materials and methods

Collection of specimens, sample sizes, and definition of grouping factors used in statistical analyses

Specimens from Northern Neotropical ariid species (Table S1) that occur in marine (m, n = 124), brackish (b, n = 88) and freshwater (f, n = 28) habitat were collected at the Caribbean coast of Venezuela and the Eastern Pacific coast of Panama (TEP) (Fig. 1A). Specimens were measured and photographed, and skulls were macerated and bleached in the field. In total, we analysed 240 specimens of *Ariopsis* (A, n = 15), *Bagre* (B, n = 50), *Cathorops* (C, n = 70), *Notarius* (N, n = 35), and *Sciades* (S, n = 70), belonging to 20 recognised and two cryptic species. An additional 19 museum specimens (n(m) = 1; n(b) = 3, n(f) = 15) of the genera *Potamarius* (two species), *Ariopsis* (two species), and *Cathorops* (one species) were included (see Table S3 for details) to extend sampling of freshwater species, that we could not cover by our own fieldwork. Note that the species identity of the specimens collected by us has been validated in a previous study using a barcoding approach [31], and a phylogenetic hypothesis based on 1768 bi-allelic single-nucleotide polymorphisms (SNPs) derived from restriction-site associated DNA-sequencing (RAD-seq) is available as well [39] and recapitulated in Fig. 1B.

3D geometric morphometric analyses

Landmarks. Landmarks (LM) are defined as homologous points on which explanations of biological processes are based upon [16]. We collected eight Type 1 landmarks (discrete juxtapositions of tissues, here of bones) plus ten Type 2 landmarks (maxima of curvature and extreme points) (Table S4) on the dorsal and ventral side of the neurocranium as illustrated in Fig. 2. Landmarks four to seven capture the shape of the mesethmoid, which is where the

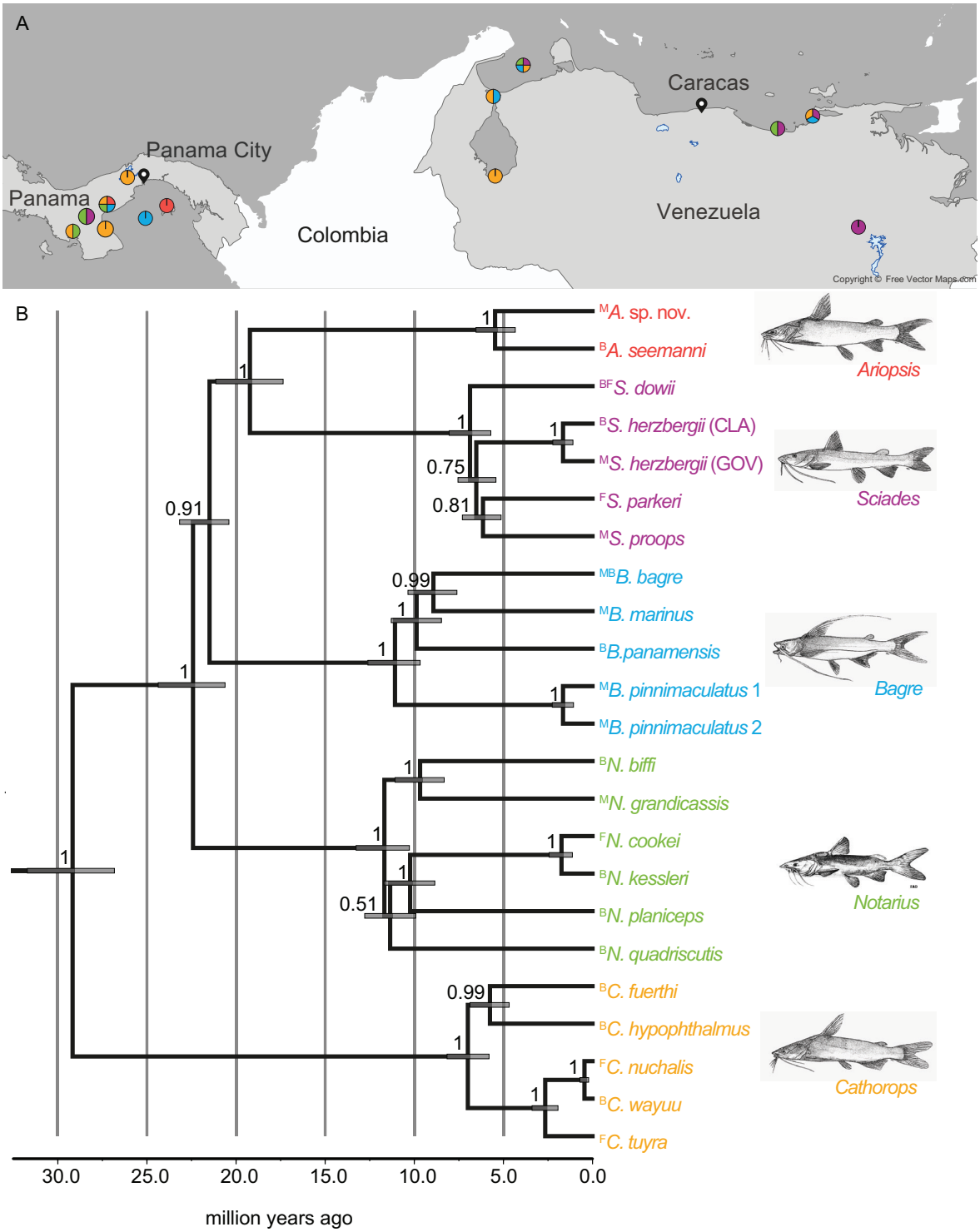


Figure 1. Geographic locations of sampling sites and phylogenetic relationships of sampled ariid species. A) Map excerpt of the Northern Neotropics focusing on Panama and Venezuela. Positions of the filled circles indicate the sampling sites, details can be found in Table S2. Sampled genera at each location are indicated by the color of the filled circles. Color key follows tip label color in panel B. B) Bayesian inference of phylogenetic relationships within Northern Neotropical Ariidae. Phylogeny modified after Stange *et al*. [39] and used for phylogenetically informed shape analyses. Bayesian posterior probability (BPP) values were calculated from 1000 trees. BPP above the line for topology. Grey bars represent 95 % height probability distribution (HPD).

maxillary teeth are attached and the olfactory and other sensory nerves exit; landmarks three and eight mark the most distal points of the ethmoid, capturing the maximal extension of the anterior part of the neurocranium (here, the vomerine teeth are attached); landmarks two and nine represent the meeting point of the sphenotic and the frontal and capture more or less the narrowest part of the neurocranium; landmarks one and ten describe the most distal points and capture the maximal extension of the posterior part of the neurocranium (or posttemporosupracleithrum); landmarks 11 and 14 outline the supraoccipital process, the exterior roof of the posterior portion of the braincase. Landmark positions were measured in the laboratory, using a MicroScribe TM G2 with an accuracy of 0.38 mm. This portable device measures coordinates in 3D space and provides x, y, and z coordinates in a text-file. To record 3D landmarks, the skulls were mounted vertically on plasticine attached to the vertebral column, allowing measurements on the ventral and dorsal side without rotating the specimen. To assess the relative measurement error, two replicates for a subset of the specimens (N = 15) were taken and analysed using Procustes ANOVA (see below). The effect of interspecific variation was larger (F = 31.40, P <0.0001) than the variation between measurements of the same specimen (F = 0.58, P = 1). Therefore, we measured each specimen once for the final geometric morphometric analysis.

Procrustes superimposition. All geometric morphometric analyses were carried out in the R package geomorph v.3.0.1 [42] and R v.3.2.4. [43] unless stated otherwise. Procrustes superimposition [44] of all measured neurocrania was performed in order to remove the effects of size, orientation and position. Missing landmarks (one to two individuals with one to two missing LMs per species set, 15 species

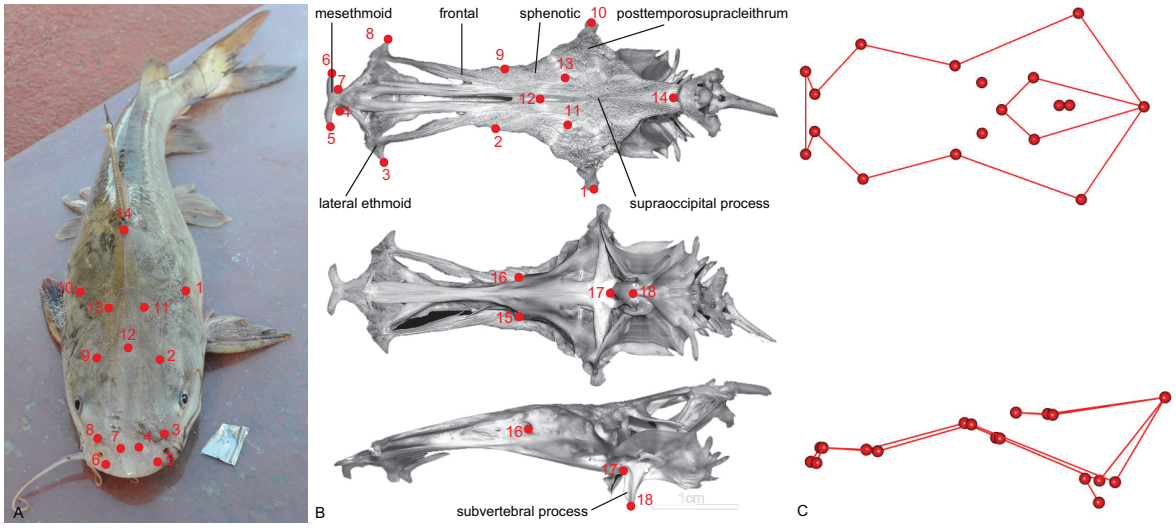


Figure 2. Positions of the eighteen landmarks that were used to study neurocranial shape change. For definitions see Table S4. A) specimen of the genus *Cathorops* with landmarks superimposed (photo credit: Madlen Stange); B) neurocranium (*Cathorops arenatus* [photo credit: Catfish Bones – The Digital Atlas Of Catfish Morphology, http://catfishbone.ansp.org/]) in dorsal, ventral and lateral view with numbered landmarks and relevant skeletal structures named; C) wireframe connecting the landmarks, that are used for illustration of shape changes.

were affected), caused by minimally broken skulls, were extrapolated using the function *estimate.missing* based on the thin-plate spline (TPS). Principal component analysis (*plotTangentSpace*) was applied in order to reduce dimensionality of the shape data and to identify major axes of variance. Morphospace plots aided to identify patterns of clustering and related skull shape changes.

Morphospace occupation. We quantified the extent of morphospace occupation of 26 recognised and two cryptic species (Tables S1 and S3) and highlighted the occupied morphospace for each genus (*Ariopsis*, *Bagre*, *Cathorops*, *Notarius*, *Potamarius*, *Sciades*) and habitat (fresh, brackish, and sea water) by convex hulls. We applied the pairwise disparity test *morphol.disparity* to calculate overall disparity (MD), the space all analysed species occupy in morphospace, by calculation of the grand mean or centroid (shape~1) in unit Procrustes variance. Further, the contribution of each genus and each habitat group to overall disparity was calculated by inferring Foote's partial disparity (PD) [25]. To do so, residuals obtained from the overall mean were used and the squared residual lengths were summed over either group mean (shape~1, groups =~genus or groups =~habitat). The resulting group-wise Procrustes variances were multiplied by number of samples per group (n) divided by total sample size (259) minus one. By this procedure the partial disparity of each group sums up to the overall disparity of the entire dataset and assertions about the percental contribution of each group to the overall disparity of all analysed specimens can be made.

To assess which groupings (genus or habitat) differ the most in terms of their relative position to each other in morphospace, a Procrustes ANOVA was used as implemented in *advanced.procD.lm*. In all analyses described above, 1000 random permutations of the residuals among groups were applied for significance testing.
Both analyses, disparity and Procrustes distance, were repeated with a subset of specimens, that included only genera with species in both habitats. This way *Bagre*, an obligate marine genus, and *Potamarius*, an obligate freshwater genus, were excluded from these analyses. We aimed at an unbiased investigation of morphospace without any specialized genera that enlarge the morphospace for one category only.

Phylogenetically informed analyses

To investigate the influence of habitat and phylogeny on skull shape variation we performed analyses that take the phylogenetic relatedness into account. The following analyses were carried out on a subset of the shape dataset (Table S1) that contains only species that are present in the available phylogeny. To this end, we first calculated mean shapes per species using the *mshape* function. The phylogenetic tree (Fig. 1 B) was taken from Stange et al. [39]. The tree is derived from a multispecies-coalescent analysis based on single-nucleotide polymorphisms and internal node calibration based on fossils instead of the common biogeographic calibra-

tion point, the final closure of the Panamanian Isthmus. It was read in using *read.nexus* (ape). The phylogeny was projected on the mean species shape scores of the tip data and the reconstructed ancestral states derived from maximum likelihood analysis using the *plotGMPhyloMorphoSpace* function. The resulting PCA plot with superimposed phylogeny facilitated visual inspection of the phylomorphospace.

Blomberg's *K* [46] and Pagel's λ [47] are estimators that assess the strength of phylogenetic signal in shape data (Procrustes coordinates) and principal components, respectively. Phylogenetic signal in this context is the association of phenotypic similarity to phylogenetic relatedness among the taxa under study. Note that Blomberg's *K* is probably the better estimator for the sample size of 22 species [48]. *K* is the ratio of the observed trait variance and the expected trait variance as predicted under Brownian motion. Therefore, *K* is small when the observed variance cannot be predicted by neutral shape evolution, and large when the shape differences are in concordance with phylogenetic distance. The estimation of *K* is implemented in the *physignal* function [49], which was run on the averaged species shape data with 1000 random permutations for significance testing. The maximum likelihood estimate of Pagel's λ was calculated for the first two principal components performed on the species means using the *pgls* function [50] in the R package Caper [51].

Rates of shape evolution

We tested whether shapes differ among habitats while accounting for phylogenetic dependencies in shape. To do so phylogenetic analyses of variance (ANOVAs) were carried out on two datasets, the more complex shape data and the derived principal components. The shape data (Procrustes coordinates) were analyzed applying a generalized least squares approach [52], as implemented in the *procD.pgls* function. The significance of differences among groups was tested in a permutation test based on residual randomization [53] (RRPP=TRUE) with 999 random permutations. The individual PCs were analysed using a simulation-based approach [54] that is implemented in *phylANOVA* in the R package phytools v.0.5-20 [55]. Significance testing was accomplished by a post-hoc t-test with Benjamini & Hochberg correction [56].

Further, we wanted to explore whether taxa that belong to a specific habitat, or taxa that diversified in the Caribbean (ancestral distribution area), or Tropical Eastern Pacific (TEP) evolved faster than their counterparts. To do so, we calculated evolutionary rates of shape evolution under the assumption of a Brownian motion model of evolution for habitat and ocean using *compare.evol.rates* [57]. Effectively, distances among taxa in morphospace are calculated after phylogenetic transformation. The test statistic is calculated from the ratio of maximum rate to minimum rate. The larger the rate ratio the bigger is the difference among rate of shape evolution among the tested groups. Significance is tested by phylogenetic simulation of tip data under Brownian motion and a common evolutionary rate for all species of the shape dataset.

Morphological and genetic distance compared

In order to test whether disparity evolves according to molecular diversity the genetic and morphological distance on species and genus level was compared based on distance matrices. Data of 22 taxa were included (Table S1), 20 of which are recognized species, the remaining two are cryptic species (contained in *Bagre pinnimaculatus* [58] and *Sciades herzbergii* (Bloch, 1794)). The genetic distance matrix was computed from 15308 single-nucleotide polymorphisms (SNPs) that were derived from restriction-site associated DNA sequencing (RAD-seq) [39]. For the computation of the distance matrix on genus level consensus sequences per genus were calculated in CodonCode Aligner v.3.7.1.1 (CodonCode, Dedham, MA, USA). The distance matrices themselves were calculated using the *dist.dna* function in the R package ape v.3.5 [59]. The morphological distance was computed using the *advanced.procD.lm* function in geomorph. The residuals were obtained from the overall mean and squared over species for yielding the species Procrustes distances, or squared over genus for the genus Procrustes distances. The species and genus matrices were compared using mantel test with Spearman rank correlation, as implemented in the R package vegan v.2.4-1 [60].

Results

Morphospace occupation, morphological disparity, and quantification of shape differences

To investigate whether species that live in a similar habitat evolved similar phenotypes – here assessed by skull shape – we inferred the shape space that is occupied by 22 Neotropical ariid species (259 individuals, 6 genera) living in marine, brackish and freshwater habitats (Fig. 3 A). The first two major axes of variance describe 60.06% of the observed overall shape variance (Fig. 3 B). Visual inspection revealed that genera mostly overlap in shape space to a certain degree and that only *Bagre* in its full shape range, and *Cathorops* and *Sciades* to the largest degree, occupy individual shape spaces. We did not detect any habitat-specific shape spaces, although freshwater species exhibit a restricted area in morphospace compared to marine and brackish species. PCs of lower ordination did not reveal differentiation of habitat shape spaces (data not shown).

The overall disparity (MD), i.e. the shape (Procrustes) variance of all analysed specimens, is 0.011. The partial disparity (PD) that each genus contributes to this total variance is smallest in Ariopsis (0.0005, 4.2% of the total variance), and largest in *Bagre* (0.0037, 33.6%). The genus-specific absolute variance was smallest in *Ariopsis* and largest in *Potamarius* (Table 1). The difference in PD and absolute variance is that the residuals of the latter were obtained from the group mean and not from the overall

mean. Therefore, absolute variance informs on the expansion in morphospace unrelated to the other genera and sample size. The distance in morphospace is largest between *Bagre* and *Potamarius* with 0.198 Procrustes distances (effect size Z = 4.94); note that *Bagre* and *Potamarius* also use the extreme margins along PC 1 (Fig. 3 A). In contrast, *Ariopsis* shares a common morphospace in the centre of the PC plot with *Sciades*, *Notarius*, and *Potamarius*.

The PD of species that live in a specific habitat contribute to MD is highest for marine and lowest for freshwater species. Marine and

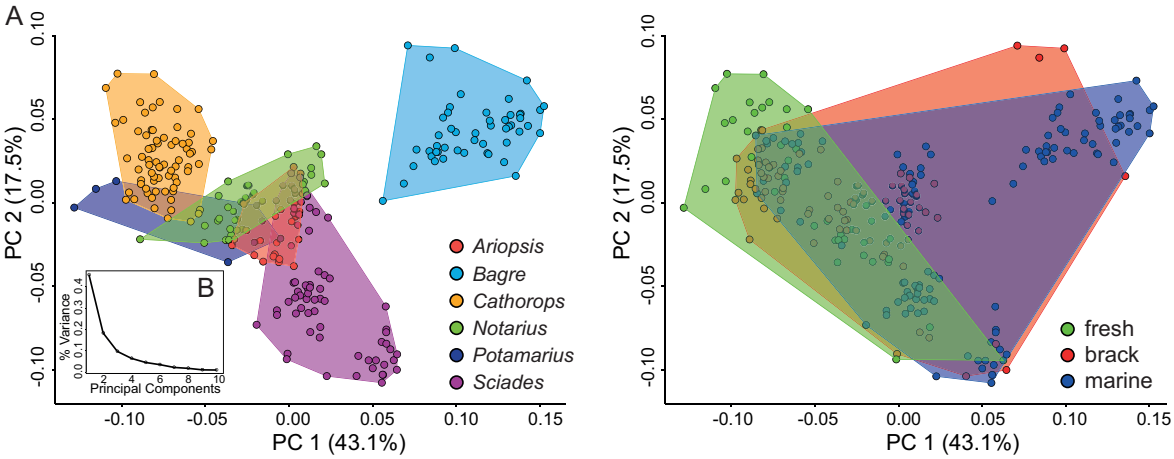


Figure 3. Morphospace occupation of sampled ariid genera. A) Morphospace of the 259 analysed ariid individuals (28 species) along principal component axes one and two. Left panel illustrates morphospace usage by genus, right panel illustrates occupancy by habitat. Convex hulls highlight the margins each group occupies in morphospace. B) Screeplot of principal components. The biggest variance is covered by the first two PCs, which account for 60.6% of the observed variation.

Table 1. Procrustes variance and Procrustes distances for genera.

	Ariopsis	***Bagre***	***Cathorops***	***Notarius***	***Potamarius***	***Sciades***
Procrustes variance	0.002619	0.004829	0.002915	0.004904	0.006997	0.005364
Partial disparity	0.000465	0.003707	0.002844	0.001044	0.000584	0.002395
	4.21 %	33.58 %	25.77 %	9.45 %	5.29 %	21.70 %
Procrustes distance						
Ariopsis	0					
Bagre	0.140786	0				
Cathorops	0.105279	0.188883	0			
Notarius	0.083067	0.149978	0.085423	0		
Potamarius	0.098709	**0.198322**	0.119722	0.106900	0	
Sciades	**0.063574**	0.134819	0.129309	0.088253	0.128617	0

P < 0.002, highest and lowest distance are highlighted in bold

Table 2. Procrustes variance and Procrustes distances for habitat groups.

	fresh	**brackish**	**marine**
Procrustes variances	0.009127	0.007463	0.010751
Partial Disparity	0.002089	0.002950	0.006000
	18.92 %	26.73 %	54.35 %
Procrustes distance			
fresh	0		
brackish	0.041000	0	
marine	0.100322	0.069159	0

$P < 0.009$

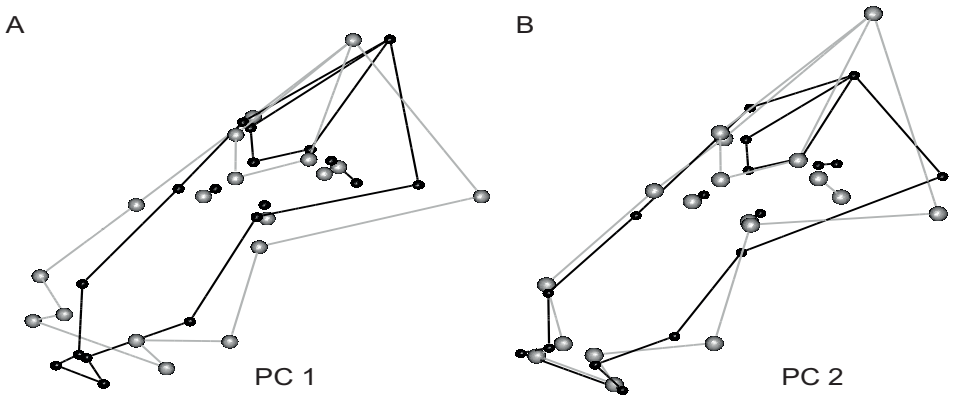


Figure 4. Shape changes along the major axes of variance. A point illustration with lines connecting landmarks (for your guidance see Fig. 2). A) Shape change along PC 1 from negative in black, to positive in gray. The neurocranium shortens, becomes broader and deeper. B) Shape change along PC 2 from negative in black, to positive in gray. The proportions of the anterior and posterior part of the neurocranium change to a shorter and broader frontal part and a larger and deeper braincase.

freshwater species have a similar group-specific variance but have the greatest distance from each other (0.10 Procrustes distances, Z = 5.25) in morphospace (summarized in Table 2, Fig. 3 A). This pattern does not change when the genera with obligate marine or freshwater species (*Bagre*, *Potamarius*) were excluded from the analyses (Table 3).

The shape changes that are associated with the two major axes of variance (Fig. 4) can be characterized as an overall deepening, broadening, and shortening of the skull along PC 1 and a shift in the proportions of anterior

Table 3. Procrustes variance and Procrustes distances for habitat groups based on genera with species in different habitats (*Ariopsis*, *Cathorops*, *Sciades*, *Notarius*). *Potamarius* and *Bagre* were excluded as they are restricted to a specific habitat.

	fresh	**brackish**	**marine**
Procrustes variances	0.007630	0.006408	0.007776
Partial Disparity	0.001256	0.002229	0.002660
	20.44 %	36.27 %	43.28 %
Procrustes distance			
fresh	0		
brackish	0.033514	0	
marine	0.069016	0.043222	0

$P < 0.02$

and posterior part of the skull accompanied by a broadening and deepening of the skull along PC 2. This means that along PC 1 (Fig. 4 A), from minus to plus, the lateral ethmoid and mesethmoid become broader, but shorter, the lateral ethmoid reshapes from being in the same plane as the frontals to being downward curved; the proportion of the frontal part, relative to the dorsal part becomes slightly smaller, because the bones of the anterior part as the lateral ethmoid, mesethmoid, and frontals reduce more in length compared to the supraoccipital process; the subvertebral process transforms from a distinct projection to forming only a small buldge; the posttemporalsupracleithrum reshapes from being flat to being downward curved resulting in an overall deeper braincase.

Along PC 2 (Fig. 4 B) the following changes from negative to positive occur: the distance between the lateral ethmoids and mesethmoid narrows; the anterior part reduces in length and the posterior part increases in length so that the proportions of the two revert along the second PC. To a lesser extent, this results from a shortening of the frontals and in a bigger extend from the elongation of the supraoccipital process.

Phylogenetically informed analyses reveal strong phylogenetic interdependence in neurocranial shape data

Estimators of phylogenetic signal find strong phylogenetic signal in the overall shape data (K= 0.6854, $P = 0.001$) and in the major axes of variance (PC 1: $\lambda = 1$, PC 2: $\lambda = 0.81$, in both cases upper bound is not significantly different from 1, $P = 1$ and $P = 0.14$, respectively). Phylogenetic ANOVAs on Procrustes coordinates and on the first two PC axes did not

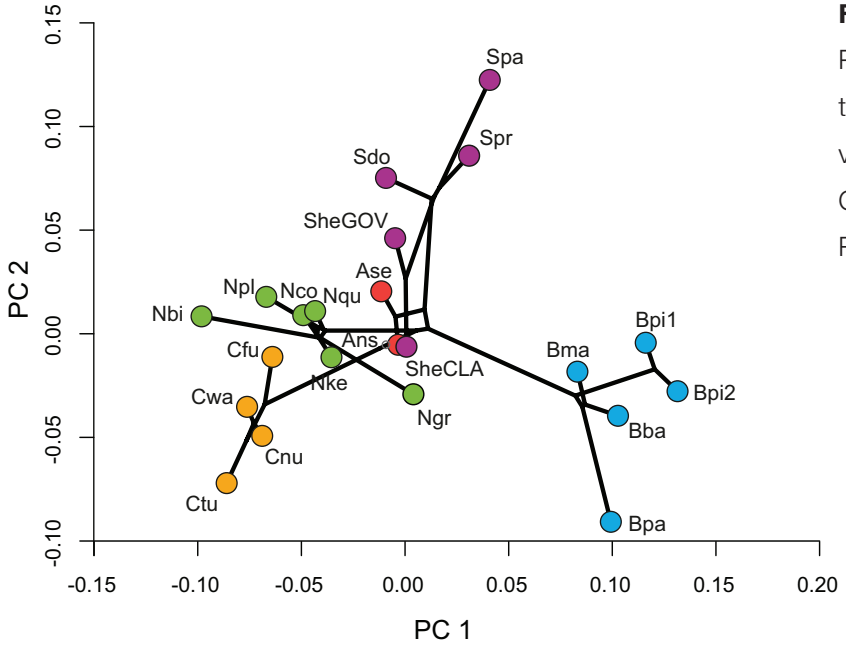


Figure 5.
Phylomorphopsace plot from the first two major axes of variance derived from 3D GM skull shape data and the RAD-seq derived phylogeny.

reveal significant differences among shapes from different habitats. However, phylogenetic ANOVA performed on PC 1 showed that freshwater and brackish species are more dissimilar compared to marine species than freshwater fish are compared to brackish species (PC1 – f/b vs. m: -2.12/-2.26, $P = 0.1$; b vs f: 0.08 $P = 0.9$). This result is in agreement with the calculation of Procrustes distances without correction for phylogenetic dependence, although non-significant.

The phylomorphospace reconstruction (Fig. 5) demonstrates little evidence for convergent evolution of species from similar habitats, as freshwater species do not evolve similar shapes. They do not appear in close proximity to each other and lines rarely cross. The comparison of molecular and morphological distance matrices of 22 species fails to reject the null hypothesis of no correlation between molecular and morphological distance on genus level ($r = 0.3091$, $P = 0.225$), but reveals correlation on species level ($r = 0.3514$, $P = 0.001$). Also, Mantel tests performed on each genus separately (except for Ariopsis, $N < 3$) yield no correlation between molecular and morphological distances among species ($P = 0.091 – 0.659$).

Rates of shape evolution are similar for marine and freshwater species

We tested whether a) Tropical Eastern Pacific (TEP) species have an elevated rate of shape evolution due to higher ocean productivity of the TEP compared to the Caribbean after the closure of the Panamanian isthmus; b) freshwater species express a lower rate of shape evolution due to the younger age of derived freshwater species, as well as increased competition in occupied niches. Calculations of rates of shape evolution revealed no signif-

icant differences of rate of shape evolution among Caribbean and TEP species, or among species from the different habitats.

Shape differentiation of cryptic species, and species and populations from different habitats

Our sample contains two so far unrecognized species, one in *B. pinnimaculatus* (called 1 and 2) from Panama and one in *S. herzbergii* from Venezuela (Clarines, brackish: CLA; Gulf of Venezuela, marine: GOV). Both cryptic species exhibit mitochondrial [31] and genomic differentiation [39] compared with their conspecifics. *B. pinnimaculatus* 1 and *B. pinnimaculatus* 2 overlap only slightly in skull shape space, the mean shape of both species clearly being different (Figs. 5 and S1), possibly caused by missing hyperossifications of the frontals in *B. pinnimaculatus* 2 [31]. Additionally to molecular divergence (Stange et al., 2016, Stange et al., 2017), *S. herzbergii* GOV and *S. herzbergii* CLA cluster apart in morphospace (Figs. 5 and S1). We also measured few *S. dowii* individuals from riverine (2 specimens) and brackish (3 specimens) populations in Panama. They showed no molecular differentiation on mitochondrial [31] or genomic level [39] and no differentiation in morphospace (Fig. 3 and S1). On the other hand, the sister taxa *C. wayuu* (brackish) and *C. nuchalis* (freshwater) share a common morphospace (Figs. 5 and S1).

Discussion

Morphospace overlap complicates delimitation of *Ariopsis* and *Sciades*

We hypothesized that neurocranium shape is to some extent impacted by the habitat that a species is living in, due to the combined effects of adaptations to the novel environment comprising new osmotic conditions, food sources, and predators. Such effects have been observed in many different vertebrates, including fishes and mammals [24–26, 61, 62].

In northern Neotropical Ariidae we found a strong phylogenetic interdependence of shapes, irrespective of their habitat of origin. The basal species (*Cathorops*) is situated at the edge of the morphospace (Fig. 3). Our results suggest that specimens can be identified via neurocranial features on genus or species level based on its position in shape space. Skull shape seems to recapitulate phylogeny, verifying the application of both molecular- and anatomy-based taxonomy. Although, anatomy-based taxonomy is restricted to genera that do not overlap in morphospace, molecular phylogenetics have a bigger potential to resolve relationships among similar looking taxa.

We found complete overlap of *Ariopsis* with part of *Potamarius*, *Sciades*, and *Notarius* morphospace. Due to shared anatomical features determined by cladistics, it has previously been argued that *Ariopsis* and *Sciades* should be grouped into a single genus [63]. Molecular studies, on the other hand, suggested

their separation [64–66]. We here provide evidence that morphological overlap in neurocranial shape – with the skull being the most prominent osteological character used in anatomy-based taxonomy–, might result in difficulties to delimit the two. Traditional morphometrics based on discrete osteological characters atomizes shape and potentially loses the information that is contained in the complex shape, which in contrast geometric morphometrics is able to capture. A previous study showed that both techniques yield comparable results [67]. Our study is an excellent example of complementation of two methods that are usually not brought together. In this specific case, geometric morphometrics yielded an advantage by projection of the genera shape space into the multidimensional shape space.

Morphospace analysis of genera with mixed habitat composition

The exclusion of the obligate marine genus *Bagre* and the obligate freshwater genus *Potamarius* enabled an unbiased view on morphospace that is occupied by genera with species in all three habitats. Marine and freshwater species retained similar variances but the PD of marine species reduced due to the exclusion of *Bagre*. This demonstrates that the general pattern of marine species occupying the largest morphospace, although insignificantly different from freshwater species, holds true. The restriction in the ability to expand in morphospace seems to lie rather on the brackish than freshwater species. This might indicate a constraint that is put on species that cope with two environmental regimes.

Phylogenetic lability of skull shape within northern Neotropical ariid genera

Concordance analyses revealed an agreement between genetic and morphological distance at the species level. Yet, the morphospace analysis of 22 Neotropical ariid species demonstrated a wide spectrum of shape evolution in a phylogenetic context considered within each genus. We found species that are young sister species and share a common morphospace (*C. wayuu* and *C. nuchalis*), species that are young sister species and diverge strongly in shape space (cryptic species in *S. herzbergii* and *B. pinnimaculatus*), species that belong to the same genus and diverged in two different oceans but occupy the same morphospace (*N. cookei* and *N. quadriscutis*) whereas they cluster far apart in shape space from their nearest relative (*N. cookei* and *N. kessleri*)

Performing Mantel tests on species that belong to a particular genus fails to find correlation of genetic and morphological distance. This demonstrates that below genus-level morphospace occupation is less influenced by phylogenetic dependence, but its maximal extension in shape space remains restricted by its hierarchical affinity to genus level.

Conclusions

The combination of 3D geometric morphometrics with a solid phylogenetic hypothesis for northern Neotropical Ariidae aided to identify patterns of skull shape diversification. We found that skull shape is mostly determined by phylogeny. Within genera, however, this constraint is weakened and skull shape might evolve under an adaptational regime as we could show that genetic divergence does not necessarily result in significant shape differences and vice versa. To test this further, studies on shape evolution should be conducted on populations-level. Freshwater species occupy the smallest place in shape space compared to brackish and marine species and differ most from marine species, possibly caused by their young clade age or competition in their new habitats. We were able to elucidate genus-specific inconsistencies regarding *Ariopsis* and *Bagre* that could not be solved by cladistic methods alone.

Funding

MS was funded by a Forschungskredit from the University of Zurich (FK-15-092); MRSV and WS were supported by the Swiss National Science Foundation Sinergia (Sinergia Grant CRSII3_136293). WS was further supported by a grant from the European Research Council (CoG "CICHLID~X").

Acknowledgements

We thank T. Barros, R. Cooke, J. D. Carrillo-Briceño for field assistance and collaboration with the officials. A. Margvelashvili, J. Weissmann, C. Zollikofer (UZH) provided access to the microscribe. D. Nelson provided museum specimens.

References

1. Arif S, Aguirre WE, Bell MA.: Evolutionary diversification of opercle shape in Cook Inlet threespine stickleback. Biol J Linn Soc 2009, 97:832–844.
2. Kimmel CB, Ullmann B, Walker C, Wilson C, Currey M, Phillips PC, Bell MA, Postlethwait JH, Cresko WA: Evolution and development of facial bone morphology in threespine sticklebacks. Proc Natl Acad Sci U S A 2005, 102:5791–5796.
3. Kimmel CB, Aguirre WE, Ullmann B, Currez M, Cresko WA: Allometric change accompanies opercular shape evolution in Alaskan threespine sticklebacks. Behaviour 2008, 145:669–691.
4. Kocher TD, Conroy JA, McKaye KR, Stauffer JR: Similar morphologies of cichlid fish in Lakes Tanganyika and Malawi are due to convergence. Mol Phylogenet Evol 1993, 2:158–165.

5. Torres-Dowdall J, Pierotti MER, Härer A, Karagic N, Woltering JM, Henning F, Elmer KR, Meyer A: Rapid and parallel adaptive evolution of the visual system of Neotropical Midas cichlid fishes. Mol Biol Evol 2017:msx143. doi: 10.1093/molbev/msx143
6. Armbruster JW, Stout CC, Hayes MM: An empirical test for convergence using African barbs (Cypriniformes: Cyprinidae). Evol Ecol 2016, 30:435–450.
7. Klingenberg CP: Morphological Integration and Developmental Modularity. Annu Rev Ecol Evol Syst 2008, 39:115–132.
8. Goswami A, Smaers JB, Soligo C, Polly PD: The macroevolutionary consequences of phenotypic integration: from development to deep time. Philos Trans R Soc B Biol Sci 2014, 369:20130254.
9. Hu Y, Ghigliotti L, Vacchi M, Pisano E, Detrich HW, Albertson RC: Evolution in an extreme environment: developmental biases and phenotypic integration in the adaptive radiation of antarctic notothenioids. BMC Evol Biol 2016, 16:142.
10. Evans KM, Waltz B, Tagliacollo V, Chakrabarty P, Albert JS: Why the short face? Developmental disintegration of the neurocranium drives convergent evolution in neotropical electric fishes. Ecol Evol 2017, 7:1783–1801.
11. Evans KM, Waltz BT, Tagliacollo VA, Sidlauskas BL, Albert JS: Fluctuations in Evolutionary Integration Allow for Big Brains and Disparate Faces. Sci Rep 2017, 7:40431.
12. Jamniczky HA, Harper EE, Garner R, Cresko WA, Wainwright PC, Hallgrímsson B, Kimmel CB: Association between integration structure and functional evolution in the opercular four-bar apparatus of the threespine stickleback, Gasterosteus aculeatus (Pisces: Gasterosteidae). Biol J Linn Soc 2014, 111:375–390.
13. Wainwright PC, Bellwood DR, Westneat MW, Grubich JR, Hoey AS: A functional morphospace for the skull of labrid fishes: Patterns of diversity in a complex biomechanical system. Biol J Linn Soc 2004, 82:1–25.
14. Westneat MW, Alfaro ME, Wainwright PC, Bellwood DR, Grubich JR, Fessler JL, Clements KD, Smith LL: Local phylogenetic divergence and global evolutionary convergence of skull function in reef fishes of the family Labridae. Proc R Soc B Biol Sci 2005, 272:993–1000.
15. Cooper WJ, Westneat MW: Form and function of damselfish skulls: rapid and repeated evolution into a limited number of trophic niches. BMC Evol Biol 2009, 9:24.
16. McCord CL, Westneat MW: Evolutionary patterns of shape and functional diversification in the skull and jaw musculature of triggerfishes (Teleostei: Balistidae). J Morphol 2016, 277:737–752.
17. Frédérich B, Sorenson L, Santini F, Slater GJ, Alfaro ME: Iterative ecological radiation and convergence during the evolutionary history of damselfishes (Pomacentridae). Am Nat 2013, 181:94–113.

18. Cooper WJ, Wernle J, Mann K, Albertson RC: Functional and Genetic Integration in the Skulls of Lake Malawi Cichlids. Evol Biol 2011, 38:316–334.
19. Price SA, Wainwright PC, Bellwood DR, Kazancioglu E, Collar DC, Near TJ: Functional Innovations and morphological diversification in parrotfish. Evolution 2010, 64:3057–3068.
20. Caldecutt WJ, Adams DC: Morphometrics of Trophic Osteology in the Threespine Stickleback, Gasterosteus aculeatus. Copeia 1998, 1998:827–838.
21. Sidlauskas BL: Testing for unequal rates of morphological diversification in the absence of a detailed phylogeny: A case study from characiform fishes. Evolution 2007, 61:299–316.
22. Sidlauskas BL: Continuous and arrested morphological diversification in sister clades of characiform fishes: A phylomorphospace approach. Evolution 2008, 62:3135–3156.
23. Foth C, Rabi M, Joyce WG: Skull shape variation in recent and fossil Testudinata and its relation to habitat and feeding ecology. Acta Zool 2016, 0:1–16.
24. Metzger KA, Herrel A: Correlations between lizard cranial shape and diet: A quantitative, phylogenetically informed analysis. Biol J Linn Soc 2005, 86:433–466.
25. Klaczko J, Sherratt E, Setz EZF: Are diet preferences associated to skulls shape diversification in xenodontine snakes? PLoS One 2016, 11:1–12.
26. Caumul R, Polly PD: Phylogenetic and environmental components of morphological variation: skull, mandible, and molar shape in marmots (Marmota, Rodentia). Evolution 2005, 59:2460–2472.
27. Diogo R: Etho-Eco-Morphological Mismatches, an Overlooked Phenomenon in Ecology, Evolution and Evo-Devo That Supports ONCE (Organic Nonoptimal Constrained Evolution) and the Key Evolutionary Role of Organismal Behavior. Front Ecol Evol 2017, 5:1–20.
28. Klingenberg CP, Marugán-Lobón J: Evolutionary covariation in geometric morphometric data: Analyzing integration, modularity, and allometry in a phylogenetic context. Syst Biol 2013, 62:591–610.
29. Zelditch ML, Ye J, Mitchell JS, Swiderski DL: Rare ecomorphological convergence on a complex adaptive landscape: body size and diet mediate evolution of jaw shape in squirrels (Sciuridae). Evolution 2017.
30. Muschick M, Nosil P, Roesti M, Dittmann MT, Harmon L, Salzburger W: Testing the stages model in the adaptive radiation of cichlid fishes in East African Lake Tanganyika. Proc R Soc B 2014, 281:20140605.
31. Stange M, Aguirre-Fernandez G, Cooke RG, Barros T, Salzburger W, Sánchez-Villagra MR: Evolution of opercle bone shape along a macrohabitat gradient: species identification using mtDNA and geometric morphometric analyses in neotropical sea catfishes (Ariidae). Ecol Evol 2016, 6:5817–5830.

32. Wilson LAB, Colombo M, Sánchez-Villagra MR, Salzburger W: Evolution of opercle shape in cichlid fishes from Lake Tanganyika - adaptive trait interactions in extant and extinct species flocks. Sci Rep 2015, 5(October):16909.
33. Colombo M, Indermaur A, Meyer BS, Salzburger W: Habitat use and its implications to functional morphology: niche partitioning and the evolution of locomotory morphology in Lake Tanganyikan cichlids (Perciformes: Cichlidae). Biol J Linn Soc 2016, 118:536–550.
34. Martin CH, Wainwright PC: Trophic novelty is linked to exceptional rates of morphological diversification in two adaptive radiations of cyprinodon pupfish. Evolution 2011, 65:2197–2212.
35. Claverie T, Wainwright PC: A morphospace for reef fishes: Elongation is the dominant axis of body shape evolution. PLoS One 2014, 9.
36. Marceniuk AP, Menezes N: Systematics of the family Ariidae (Ostariophysi, Siluriformes), with a redefinition of the genera. Zootaxa 2007, 1416:1–126.
37. Betancur-R. R: Molecular phylogenetics supports multiple evolutionary transitions from marine to freshwater habitats in ariid catfishes. Mol Phylogenet Evol 2010, 55:249–258.
38. Vega GC, Wiens JJ: Why are there so few fish in the sea? Proc R Soc B 2012, 279:2323–2329.
39. Stange M, Sánchez-Villagra MR, Salzburger W, Matschiner M: Bayesian Divergence-Time Estimation with Genome-Wide SNP Data of Sea Catfishes (Ariidae) Supports Miocene Closure of the Panamanian Isthmus. bioRxiv 2017. doi: http://dx.doi.org/10.1101/102129.
40. Schneider B, Schmittner A: Simulating the impact of the Panamanian seaway closure on ocean circulation, marine productivity and nutrient cycling. Earth Planet Sci Lett 2006, 246:367–380.
41. Bookstein FL: Introduction to methods for landmark data. In Prooceedings of the Michigan Morphometrics Workshop. Volume 2. Edited by Rohlf FJ, Bookstein FL. Ann Arbor, MI, MI: The University of Michigan Museum of Zoology; 1990:216–225.
42. Adams DC, Otárola-Castillo E: Geomorph: an R Package for the Collection and Analysis of Geometric Morphometric Shape Data. Methods Ecol Evol 2013, 4:393–399.
43. R Core Team: R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016. Available from: http://www.rproject.org.
44. Rohlf FJ, Slice D: Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst Biol 1990, 39:40–59.
45. Zelditch ML, Swiderski DL, Sheets HD: Geometric Morphometrics for Biologists : A Primer. 2nd edition. Chapter 10: Magnitude and Structure of Morphological Diversity, pp.282. London: Elsevier Inc.

46. Blomberg SP, Garland T, Ives AR: Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 2003, 57:717–745.
47. Pagel M: Inferring the historical patterns of biological evolution. Nature 1999, 401:877–884.
48. Kamilar JM, Cooper N: Phylogenetic signal in primate behaviour, ecology and life history. Philos Trans R Soc B Biol Sci 2013, 368:20120341.
49. Adams DC: A generalized K statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. Syst Biol 2014, 63:685–697.
50. Freckleton R, Harvey PH, Pagel M: Phylogenetic analysis and comparative data: a test and review of evidence. Am Nat 2002, 160:712–726.
51. Orme D: The caper package: comparative analysis of phylogenetics and evolution in R. R Package version 0.5, 2013:1–36.
52. Adams DC: A method for assessing phylogenetic least squares models for shape and other high-dimensional multivariate data. Evolution 2014, 68:2675–2688.
53. Collyer ML, Sekora DJ, Adams DC: A method for analysis of phenotypic change for phenotypes described by high-dimensional data. Heredity 2015, 115:357–365.
54. Garland T, Dickerman AWA, Janis CM, Jones JA, Url S: Phylogenetic analysis of covariance by computer simulation. Syst Biol 1993, 42:265–292.
55. Revell LJ: phytools: An R package for phylogenetic comparative biology (and other things). Methods Ecol Evol 2012, 3:217–223.
56. Benjamini Y, Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 1995, 57:289–300.
57. Adams DC: Quantifying and comparing phylogenetic evolutionary rates for shape and other high-dimensional phenotypic data. Syst Biol 2014, 63:166–177.
58. Steindachner F: Sitzungsberichte Der Kaiserlichen Akademie Der Wissenschaften, Mathematisch-Naturwissenschaftlichen Classe. Volume 72. Wien; 1876.
59. Paradis E, Claude J, Strimmer K: APE: analyses of phylogenetics and evolution in R language. Bioinformatics 2004, 20:289–290.
60. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H: vegan: Community Ecology Package, R package version 2.3-5. 2016.
61. Kimmel CB, Hohenlohe P a., Ullmann B, Currey M, Cresko W a.: Developmental dissociation in morphological evolution of the stickleback opercle. Evol Dev 2012, 14:326–337.
62. Wilson LAB, Colombo M, Hanel R, Salzburger W, Sánchez-Villagra MR: Ecomorphological disparity in an adaptive radiation: opercular bone shape and stable isotopes in Antarctic icefishes. Ecol Evol 2013, 3:3166–3182.

63. Marceniuk AP, Menezes NA, Britto MR: Phylogenetic analysis of the family Ariidae (Ostariophysi: Siluriformes), with a hypothesis on the monophyly and relationships of the genera. Zool J Linn Soc 2012, 165:534–669.
64. Betancur-R. R, Acero P. A, Bermingham E, Cooke R: Systematics and biogeography of New World sea catfishes (Siluriformes: Ariidae) as inferred from mitochondrial, nuclear, and morphological evidence. Mol Phylogenet Evol 2007, 45:339–357.
65. Betancur-R. R: Molecular phylogenetics and evolutionary history of ariid catfishes revisited: a comprehensive sampling. BMC Evol Biol 2009, 9:175.
66. Betancur-R. R, Orti G, Stein AM, Marceniuk AP, Pyron RA: Apparent signal of competition limiting diversification after ecological transitions from marine to freshwater habitats. Ecol Lett 2012, 15:822–830.
67. Hetherington AJ, Sherratt E, Ruta M, Wilkinson M, Deline B, Donoghue PCJ: Do cladistic and morphometric data capture common patterns of morphological disparity? Palaeontology 2015, 58:393–399.

Chapter 3

Phylomorphospace and morphological diversity of northern Neotropical Ariidae reveals conservatism among genera despite habitat transitions

Supplementary material

Table S1. List of species and number of specimens used for morphospace analysis.

Species	Number of specimens
Ariopsis seemanni (Günther 1864)	5
Ariopsis jimenezi Marceniuk et al., (in review)	10
Bagre bagre (Linnaeus, 1766)	5
Bagre marinus (Mitchill, 1815)	17
Bagre panamensis (Gill 1863)	3
Bagre pinnimaculatus (Steindachner 1876), called *B. pinnimaculatus* 1 in text	4
Bagre aff. pinnimaculatus (undescribed, reported in Stange *et al.*, 2016 and Stange *et al.*, 2017), called *B. pinnimaculatus* 2 in text	21
Cathorops fuerthii (Steindachner, 1876)	5
Cathorops nuchalis (Günther 1864)	11
Cathorops steindachneri (Gilbert & Starks 1904)	1
Cathorops wayuu Betancur-R., Acero P. & Marceniuk (2012)	46
Cathorops tuyra (Meek & Hildebrand, 1923)	7
Notarius biffi Betancur-R and Acero (2004)	1
Notarius cookei (Acero and Betancur-R 2002)	7
Notarius grandicassis (Valenciennes 1840)	9
Notarius planiceps (Steindachner, 1877)	1
Notarius quadriscutis (Valenciennes, 1840)	7
Notarius kessleri (Steindachner 1876)	10
Sciades dowii (Gill 1863)	5
Sciades proops (Valenciennes, 1840)	19
Sciades herzbergii (Bloch, 1784), called *S. herzbergii* GOV in text	26
Sciades aff. herzbergii (undescribed, reported in Stange *et al.*, 2016 and Stange *et al.*, 2017), called *S. herzbergii* CLA in text	19
Sciades parkeri (Traill 1832)	1

Table S2. Geographic locations of sampling sites with GPS coordinates and habitat (marine, brackish, freshwater) attribution.

Location	Country	GPS coordinates	Habitat	Species
Lago de Maracaibo/ Isla de Toas	Zulia, Venezuela	10°57'9.50"N 71°38'49.54"W	brackish	*C. wayuu*
Lago de Maracaibo/ Isla de San Carlos	Zulia, Venezuela	10°59'55.1"N 71°36'19.8"W	brackish	*B. bagre*
Lago de Maracaibo/ Puerto Concha	Zulia, Venezuela	9°05'46.0"N 71°42´52"W	freshwater	*C. nuchalis*
Lago de Maracaibo/ Guarico	Zulia, Venezuela	10°43'52.0"N 71°31'40.2"W	brackish	*C. wayuu*
Gulf of Venezuela	Falcon, Venezuela	11°14'15.3"N 70°30'53.1" W	marine	*S. proops, S. herzbergii, C. wayuu, B. marinus, B. bagre, N. grandicassis*
Clarines	Anzoategui, Venezuela	10° 3'46.76"N 65°11'5.23"W	brackish	*N. quadriscutis, S. aff. herzbergii, S. proops*
Puerto La Cruz	Anzoategui, Venezuela	10°12'58.63"N 64°38'39.16"W	marine	*C. wayuu, B. marinus S. proops*
Ciudad Bolivar	Bolívar, Venezuela	8°8'51.46" N 63°32'10.68"W	freshwater	*S. parkeri*
Pearl Islands/Casaya island	Panama	8°34'38.64"N 79°3'3.636" W	marine	*A. jeminezi*
Puente del Rio Chagres	Panama	9°11'34.66"N 79°39'9.42"W	freshwater	*C. tuyra*
Rio Hato	Panama	8°20'32.4"N 80°09'56.4"W	brackish	*C. fuerthi, S. dowii*
Rio Santa Maria	Panama	8° 6'20.30"N 80°33'16.06"W	freshwater	*N. cookei, S. dowii*
Rio Parita	Panama	8°01'13.69"N 80°27'11.15"W	brackish	*C. fuerthi*
Rio Estero Salado	Panama	8°10'30.324"N 80°29'35.052" W	brackish	*B. pinnimaculatus, B. panamensis, N. planiceps, N. kessleri, A. seemanni*
Rio San Pedro	Panama	7°50' 59.208"N 81°07' 3.972" W	brackish	*N. kessleri, N. biffi*
Puerto Caimito	Panama	8°52'18.88"N 79°42'32.99"W	marine	*S. dowii*
Gulf of Panama	Panama	8°48'56.55"N 79°22'50.85"W	marine	*B. pinnimaculatus, B. aff pinnimaculatus*

Table S3. List of museum specimens used for extension of morphospace analysis.

catalog no. UMMZ	taxonomy used by UMMZ	taxonomy used in the present study	locality and habitat
196479-1	*Arius assimilis*	*Ariopsis assimilis*[a]	Mexico, Laguna Bacalur, Caribbean drainage, brackish
197184-3	*Arius assimilis*	*Ariopsis assimilis*[a]	Guatemala, Rio Dulce, freshwater
197214-1	*Arius assimilis*	*Ariopsis assimilis*[a]	Guatemala, Rio Nimblaja, freshwater
197214-2	*Arius assimilis*	*Ariopsis assimilis*[a]	Guatemala, Rio Nimblaja, freshwater
179147-1	*Arius felis*	*Ariopsis felis*[a, b]	USA, Florida, Bay Gulf of Mexico, marine
179147-2	*Arius felis*	*Ariopsis felis*[a, b]	USA, Florida, Bay Gulf of Mexico, marine
186995-2	*Arius felis*	*Ariopsis felis*[a, b]	USA, Florida, Monroe, Indian Key, marine
248826-1	*Arius felis*	*Ariopsis felis*[a, b]	USA, Florida, Osprey Nest, brackish
198711-1	*Arius sp.*	*Cathorops aguadulce*[b]	Guatemala, Laguna de Ronpiro, Usumacinta drainage, freshwater
198711-2	*Arius sp.*	*Cathorops aguadulce*[b]	Guatemala, Laguna de Ronpiro, Usumacinta drainage, freshwater
198712-1	*Arius aguadulce*	*Cathorops aguadulce*[b]	Mexico, Rio Salinas, Usumacinta drainage, freshwater
143496-1	*Potamarius nelsoni*	*Potamarius nelsoni*	Guatemala, Rio de la Pasion, Usumacinta drainage, freshwater
198713-2	*Potamarius nelsoni*	*Potamarius nelsoni*	Guatemala, Rio de la Pasion, Usumacinta drainage, freshwater
198721	*Potamarius nelsoni*	*Potamarius nelsoni*	Guatemala, Rio de la Pasion, Usumacinta drainage, freshwater
188048-1	*Potamarius usumacintae*	*Potamarius usumacintae*	Guatemala, Rio de la Pasion, Usumacinta drainage, freshwater
190074-2	*Potamarius usumacintae*	*Potamarius usumacintae*	Guatemala, Rio de la Pasion, Usumacinta drainage, freshwater
190074-1	*Potamarius usumacintae*	*Potamarius usumacintae*	Guatemala, Rio de la Pasion, Usumacinta drainage, freshwater
198716	*Potamarius usumacintae*	*Potamarius usumacintae*	Guatemala, Rio de la Pasion, Usumacinta drainage, freshwater
198715	*Potamarius usumacintae*	*Potamarius usumacintae*	Guatemala, Rio de la Pasion, Usumacinta drainage, freshwater

UMMZ: University of Michigan Museum of Zoology

[a] [34], [b] [35]

Table S4. Definition of the 18 landmarks used to study the neurocranial morphology and shape changes in ariid species from freshwater, brackish, and marine habitat.

Landmark #	defined by
1	most lateral point of left posttemporosupracleithrum
2	lateral suture of left sphenotic and frontal
3	most lateral point of left lateral ethmoid
4	medial curvature of left portion of mesethmoid
5	most lateral point of left portion of mesethmoid
6	most lateral point of right portion of mesethmoid
7	medial curvature of right portion of mesethmoid
8	most lateral point of right lateral ethmoid
9	lateral suture of right pterotic and sphenotic
10	most lateral point of right posttemporosupracleithrum
11	sutures meeting of left sphenotic, pterotic and supraoccipital
12	sutures meeting of left and right frontal meeting supraoccipital
13	sutures meeting of right sphenotic, pterotic and supraoccipital
14	caudal part of supraoccipital process
15	left opening of visual nerve
16	right opening of visual nerve
17	opening of aortic canal
18	tip of subvertebral process

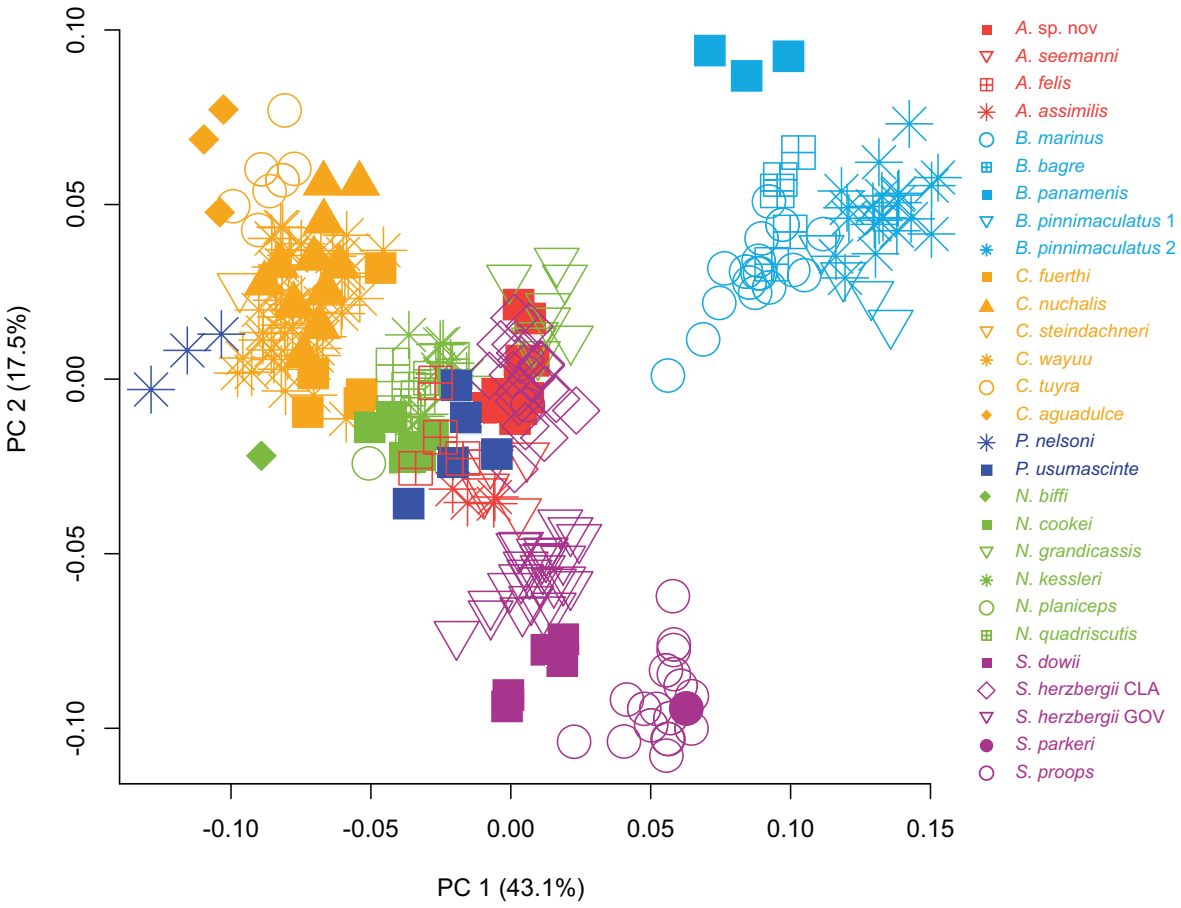


Figure S1. Morphospace analysis of 28 northern Neotropical ariid species. Genus affinity highlighted by colour, species affinity highlighted by symbols. This plot is identical to Figure 3 but the morphospace for each species is also indicated here.

Conclusion

Evolution in Northern Neotropical Catfishes – Integrating Genomics and Morphometrics

The study on Neotropical Ariidae was originally set out to mainly investigate macrohabitat transitions. In the process, the group has proven to be much more resourceful. We gained insights into the biogeography, natural history, taxonomy, and shape evolution of theses fishes.

The presented PhD thesis confirms that Neotropical Ariidae originated from an area that is the Western Atlantic today. The dating of lineage splitting events between Caribbean and Eastern Pacific ariid species provides additional biological evidence for a temporary Miocene emergence of the Panamanian isthmus.

Molecular (mitochondrial and nuclear) and shape data confirmed the presence of two new species, one closely related to *Bagre pinnimaculatus* from the Eastern Pacific side of Panama, and one closely related to *Sciades herzbergii* from Venezuela; the presence of *Notarius biffi* in Central America, outside of its reported distribution range, was verified. The shape data reveal the unique position that *Bagre* occupies in shape space, but phylogenetic analyses do not support a unique taxonomic status of that genus. The joint analysis of molecular and shape data elucidates why *Ariopsis* and *Sciades* are so difficult to distinguish using cladistics markers. The two genera share a common morphospace although having diverged since millions of years. Strong evidence based on skull shape and otolith morphology is provided for the affiliation of *Potamarius usumascinate* to the genus *Potamarius*, its membership to this genus is currently disputed.

The study of shape evolution in relation to habitat demonstrates that in this particular family, independent of the investigated feature – whether opercle bone or skull –, the phylogenetic relationship predicts well the potential position of species in morphospace. Habitat-specific opercle shapes were demonstrated to exist and confirm similar findings in previous studies in teleost fishes (Arif *et al.*, 2009; Kimmel *et al.*, 2012; Wilson *et al.*, 2013, 2015). Yet in this particular clade the phylogenetic component has a much bigger impact than habitat. The analysis of skull shape in morphospace revealed that freshwater ariids are most distinct from marine species and use a much smaller space than marine and brackish species in morphospace. Although elevated rates of morphological evolution have been found in sea catfishes (Rabosky *et al.*, 2013) we found no elevated rates of shape evolution comparing freshwater and marine species.

In general, this work adds to the understanding of the geographic occurrence and biology of the analysed species. We found species and populations at places where they were not reported before, just to name one – *Sciades parkeri* 200 km upstream in the Orinoco River, far from its native distribution area at the coasts of Guyana to Brazil. We report on freshwater populations of native marine/brackish species, which were not mentioned before in the literature or on the common fish database (fishbase.org). To extend this knowledge, I am currently preparing descriptions of seven so far undescribed ariid otoliths (ear stones), which are taxonomically very important fea-

tures for palaeontologists, as well as thin-sections of otoliths of all studied species for age determination and comparisons of growth rates among species. Ariidae offer a wealth of knowledge on very different biological topics, but it also seems like that the understanding of their natural history is key for further explorations. One cannot study a species without studying its environmental context.

References

Arif, S., Aguirre, W.E. & Bell, M.A. 2009. Evolutionary diversification of opercle shape in Cook Inlet threespine stickleback. *Biol. J. Linn. Soc.* **97**: 832–844.

Bloom, D.D., Weir, J.T., Piller, K.R. & Lovejoy, N.R. 2013. Do freshwater fishes diversify faster than marine fishes? A test using state-dependent diversification analyses and molecular phylogenetics of New World silversides (Atherinopsidae). *Evolution.* **67**: 2040–2057.

Kimmel, C.B., Hohenlohe, P.A., Ullmann, B., Currey, M. & Cresko, W.A. 2012. Developmental dissociation in morphological evolution of the stickleback opercle. *Evol. Dev.* **14**: 326–337.

Rabosky, D.L., Santini, F., Eastman, J., Smith, S.A., Sidlauskas, B., Chang, J., *et al.* 2013. Rates of speciation and morphological evolution are correlated across the largest vertebrate radiation. *Nat. Commun.* **4**: 1958.

Wilson, L.A.B., Colombo, M., Hanel, R., Salzburger, W. & Sánchez-Villagra, M.R. 2013. Ecomorphological disparity in an adaptive radiation: opercular bone shape and stable isotopes in Antarctic icefishes. *Ecol. Evol.* **3**: 3166–3182.

Wilson, L.A.B., Colombo, M., Sánchez-Villagra, M.R. & Salzburger, W. 2015. Evolution of opercle shape in cichlid fishes from Lake Tanganyika - adaptive trait interactions in extant and extinct species flocks. *Sci. Rep.* **5**: 1–15.

Author contributions

Chapter 1

MS, WS, and MRSV conceived and designed the study. MS, GA, TB, RGC and MRSV carried out the sampling of the data. MS carried out the morphometric and molecular data analysis, interpretation, and manuscript drafting. All authors contributed to the manuscript modification and finalization.

Chapter 2

MS, WS, and MRSV conceived and designed the study on Ariidae. MS, MRSV carried out the sampling of the fishes. MM designed, carried out and interpreted the simulation study and performed SNAPP analysis on the army ants and catfish data. MS carried out SNP marker generation, and interpretation of the catfish data in the background of the emergence of the Panamanian isthmus. MM and MS drafted the manuscript. All authors contributed to the manuscript modification and finalization.

Chapter 3

MS and MRSV conceived and designed the study. GA carried out collection of skull shape data, MS analysed and interpreted the shape data. MS drafted the manuscript. MS, MRSV, GA, and WS contributed to the manuscript modification and finalization.

Co-authored publication linked to this dissertation (abstract)

Manuela Fuchs, Madeleine Geiger, Madlen Stange, Marcelo R. Sánchez-Villagra (2015) Growth trajectories in the cave bear and its extant relatives: an examination of ontogenetic patterns in phylogeny. BMC Evolutionary Biology, 15:239 doi:10.1186/s12862-015-0521-z

Background

The study of postnatal ontogeny can provide insights into evolution by offering an understanding of how growth trajectories have evolved resulting in adult morphological disparity. The *Ursus* lineage is a good subject for studying cranial and mandibular shape and size variation in relation to postnatal ontogeny and phylogeny because it is at the same time not diverse but the species exhibit different feeding ecologies. Cranial and mandibular shapes of *Ursus arctos* (brown bear), *U. maritimus* (polar bear), *U. americanus* (American black bear), and the extinct *U. spelaeus* (cave bear) were examined, using a three-dimensional geometric morphometric approach. Additionally, ontogenetic series of crania and mandibles of *U. arctos* and *U. spelaeus* ranging from newborns to senile age were sampled.

Results

The distribution of specimens in morphospace allowed to distinguish species and age classes and the ontogenetic trajectories *U. arctos* and *U. spelaeus* were found to be more similar than expected by chance. Cranial shape changes during ontogeny are largely size related whereas the evolution of cranial shape disparity in this clade appears to be more influenced by dietary adaptation than by size and phylogeny. The different feeding ecologies are reflected in different cranial and mandibular shapes among species.

Conclusions

The cranial and mandibular shape disparity in the *Ursus* lineage appears to be more influenced by adaptation to diet than by size or phylogeny. In contrast, the cranial and mandibular shape changes during postnatal ontogeny in *U. arctos* and *U. spelaeus* are probably largely size related. The patterns of morphospace occupation of the cranium and the mandible in adults and through ontogeny are different.

Curriculum Vitae

Madlen Stange
University of Zurich
Palaeontological Institute and Museum
Karl-Schmid-Strasse 4
CH-8006 Zurich

Phone: +41 44 634 23 28
Mail: madlen.stange@pim.uzh.ch
Web: https://www.pim.uzh.ch/p/stange

PERSONAL INFORMATION

Name	STANGE
First name	Madlen
Date and place of birth	24.10.1985 Frankfurt/Oder
Nationality	German

EDUCATION

10/2010 – **06/2013** University of Potsdam (D)	**Master studies** in Molecular Biology Master of Science; overall grade: 1.2, "with distinction" **Master thesis** "A male color dimorphism in an African cichlid fish" in collaboration with the University of Basel (supervisors: Prof. Walter Salzburger, Prof. Ralph Tiedemann, grade: 1.0, "excellent")
10/2006 – 09/2010 University of Potsdam (D)	**Bachelor studies** in Life Sciences, focus in genome research Bachelor of Science; overall grade: 2.3, "good" **Bachelor thesis** "Genetic variability of selected Calvin cycle genes in Arabidopsis thaliana", Evolutionary Biology, University of Potsdam (supervisors: Prof. Ralph Tiedemann, Prof. Michael Lenhard, grade: 1.0, "excellent")
06/2005 Geschwister-Scholl-Gymnasium, Fürstenwalde/Spree	**allgemeine Hochschulreife (Abitur)**

ACADEMIC CAREER

01/2014 – expected 05/2017 University of Zürich (CH)	**PhD candidate** Research on the morphological and molecular evolution after habitat transition in Neotropical sea catfishes, supervised by Prof. Marcelo Sánchez, Paleontological Institute and Museum, and Prof. Walter Salzburger, University of Basel, **maternity leave** 06/2015-10/2015
07/2013 – 12/2013 University of Zürich (CH)	**Research assistant** Investigation of the possibility to conduct a research project on the habitat evolution in Neotropcial sea catfishes, with Prof. Marcelo Sánchez, Paleontological Institute and Museum
04/2012 – 10/2012 University of Basel (CH)	**Internship** Introduction to next-generation sequencing – library preparation and computational analyses at the Evolutionary Biology unit (Prof. Salzburger), Zoological Institute, Department Environmental Sciences
11/2010 – 03/2013 University of Potsdam (D)	Core Group member of the *Evolution Across Scales* Initiative funded by the Volkswagen foundation
09/2007 – 03/2012 University of Potsdam (D)	**Research assistant** Working on the genetic variability in photosynthesis genes in accessions of *Arabidopsis thaliana*, in the framework of *Go*FORSYS, Evolutionary Biology/Systematic Zoology (Prof. Tiedemann), Institute of Biochemistry and Biology

RESEARCH ARTICLES

Madlen Stange, Gabriel Aguirre-Fernández, Richard G. Cooke, Tito Barros, Walter Salzburger, Marcelo R. Sánchez-Villagra (2016) Evolution of opercle bone shape along a macrohabitat gradient: species identification using mtDNA and geometric morphometric analyses in neotropical sea catfishes (Ariidae). *Ecology and Evolution*, 6(16):5817-5830.

Manuela Fuchs, Madeleine Geiger, **Madlen Stange**, Marcelo R. Sánchez-Villagra (2015) Growth trajectories in the cave bear and its extant relatives: an examination of ontogenetic patterns in phylogeny. *BMC Evolutionary Biology*, 15:239.

Catarina S Mateus*, **Madlen Stange***, Daniel Berner, Marius Rösti, Bernardo R Quintella, M Judite Alves, Pedro R Almeida, Walter Salzburger (2013) Rad-seq reveals strong genetic divergence between sympatric European river and brook lampreys, *Current Biology* 23(15):R649-R650. (* shared first authorship)

RESEARCH GRANTS AND SCHOLARSHIPS

11/2015 – 10/2016	Forschungskredit University of Zürich (55.821 CHF), salary
10/2013 – 11/2013	Exchange visit grant, European Science Foundation (3.500 EUR), for an international collaboration with Dr. Marta Barluenga, National Museum of Natural Sciences, Dept. Biodiversity and Evolutionary Biology, Madrid, Spain
05/2012 – 10/2012	Scholarship for a master thesis abroad (PROMOS), German Academic Exchange Service (1.125 EUR), for an external thesis at the University of Basel, Switzerland

POSTER PRESENTATIONS AND TALKS

02/2017 Bern, CH	Talk, *Bayesian divergence time estimation and skull shape evolution in sea catfishes (Ariidae)*, Biology17 (awarded with the Best Talk Award in the section Bioinformatics and Phylogenomics)
09/2016 Kiel, Germany	Poster, **Madlen Stange** (presenter), Michael Matschiner, Gabriel Aguirre-Fernández, Walter Salzburger, Marcelo R. Sánchez-Villagra. *Morphological Evolution in Neotropical Ariidae*. 109th Annual Meeting of the German Zoological Society (DZG)
06/2016 Austin, TX, USA	Talk, *Morphological evolution in Neotropical Ariidae.* Evolution 2016
08/2015 Lausanne, CH	Poster, **Madlen Stange**, Gabriel Aguirre-Fernández, Walter Salzburger, Marcelo Sánchez (presenter). *Morphological adaptation in a marine-freshwater habitat transition of Northern Neotropical catfishes*. 15th Congress of the European Society for Evolutionary Biology
02/2015 eawag Dübendorf, CH	Poster, **Madlen Stange** (presenter), Gabriel Aguirre-Fernández, Jorge Carillo, Walter Salzburger, Marcelo Sánchez. *Evolution in Northern Neotropical Catfishes - A Molecular and Morphological Approach*. Biology15
10/2014 Panama-City, Panama	Talk, *Evolution in Northern Neotropical Catfishes – A Molecular and Morphological Approach*. Smithsonian Tropical Research Institute
09/2012 Leuven, Belgium	Poster, **Madlen Stange** (presenter), Walter Salzburger. *Genetic basis of a male color variation in Cyprichromis leptosoma*. Cichlid Science
06/2012 Dublin, Ireland	Poster, **Madlen Stange** (presenter), Sandra Schwarte, Fanny Wegner, Ralph Tiedemann. *Genetic variability of Calvin cycle genes in Arabidopsis accessions*.

	Annual Conference of the Society for Molecular Biology & Evolution
08/2011 Tübingen, Germany	Poster, **Madlen Stange** (presenter), Sandra Schwarte, Fanny Wegner, Ralph Tiedemann. *Genetic variability of Calvin cycle genes in accessions of Arabidopsis thaliana*. 13th Congress of the European Society for Evolutionary Biology

OUTREACH

12/2016 Paläontologisches Museum der Universität Zürich	Talk, *Von Knochen und Genen mariner Welse*. Öffentliche Museumsvorträge HS16
09/2014 Paläontologisches Museum der Universität Zürich	Talk, *Welse in Südamerika - urtümliche Fische aus Venezuela*. Lange Nacht der Zürcher Museen

FIELDTRIPS

10/2014 Panama	Collection of DNA and bone samples from catfishes in the framework of my PhD project, based at the Smithsonian Tropical Research Institute (STRI)
07/2014 Venezuela	Collection of DNA and bone samples from catfishes in the framework of my PhD project, based at Universidad del Zulia, Maracaibo
11/2013 Venezuela	Collection of DNA and bone samples from catfishes in preparation for my PhD project
08/2012 Zambia	Collection of tissue samples for DNA and RNA extraction at Lake Tanganyika in the framework of my master thesis

TEACHING EXPERIENCE

2014 – 2017	Teaching the introductory course to geometric morphometrics for Bachelor students (1/2 day) and supervision of student projects during Blockkurs (3.5 w, BIO262), University of Zürich
2014-2015	Co-supervision of Master student Manuela Fuchs, University of Zürich

MEDIA

derStandard (11.8.2013). Forscher lösen altes Rätsel um urtümliche Neunaugen.

Neue Zürcher Zeitung (7.8.2013): Urtümliche Neunaugen sind zwei Arten.

Tages-Anzeiger (7.8.2013): Genau angeschaute Neunaugen.

Spiegel online (7.8.2013): Evolution: Biologen lösen Neunaugen-Rätsel.

Bauernzeitung (7.8.2013): Genetisch bestätigt: Es gibt zwei Arten von Neunaugen.

20 Minuten (7.8.2013): Neunaugen: Rätsel gelöst.

Blick am Abend (6.8.2013): Basler Forscher finden den Unterschied.

Uni News (University of Basel, 6.8.2013): Basler Zoologen klären altes Rätsel um Neunaugen auf: Es sind zwei Arten.

Zürich, 13. March 2017